TURING 图灵电子与电气工程丛书

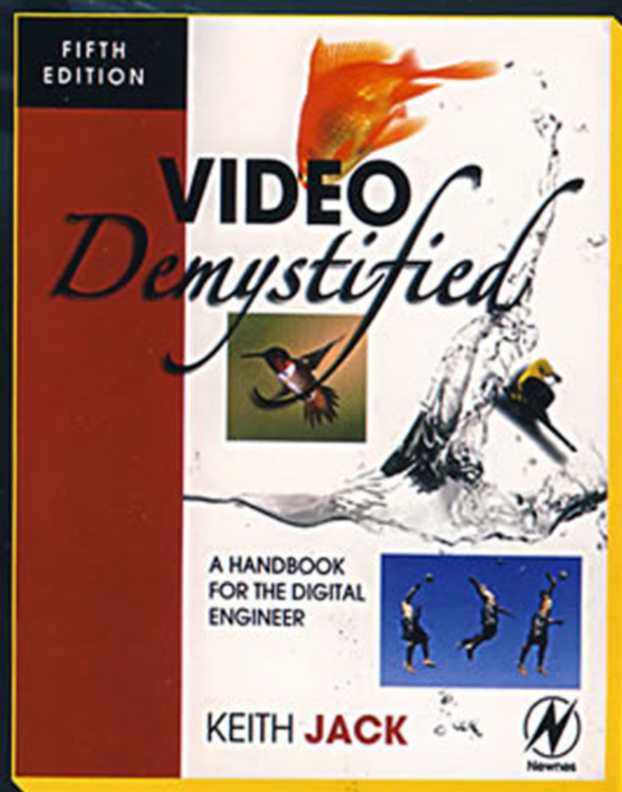
ELSEVIER

（第5版）

视频技术手册

Video Demystified
(Fifth Edition)

[美] Keith Jack 著
杨 征 田尊华
张杰良 陈 虎 等译





人民邮电出版社
POSTS & TELECOM PRESS

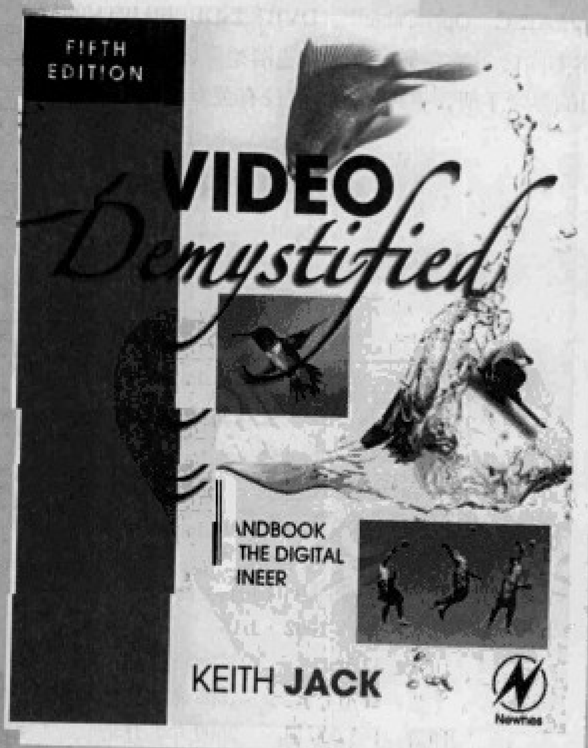


图灵电子与电气工程丛书

视频技术手册（第5版）

Video Demystified (Fifth Edition)

[美] Keith Jack 著
杨 征 田尊华
张杰良 陈 虎



人民邮电出版社
北 京

图书在版编目（CIP）数据

视频技术手册：第5版/（美）杰克（Jack, K.）著；杨征等译. —北京：人民邮电出版社，2009.8
（图灵电子与电气工程丛书）
书名原文：Video Demystified, Fifth Edition
ISBN 978-7-115-21020-3

I. 视… II. ①杰… ②杨… III. 视频系统－技术手册
IV. TN94-62

中国版本图书馆CIP数据核字（2009）第097229号

内容提要

本书较为全面地介绍了最新的数字视频相关技术，共分三个部分。第一部分介绍视频格式、颜色空间、视频信号、视频接口、数字视频处理需求和彩色电视制式标准等基本概念和基本原理。第二部分论述各种视频压缩标准，包括H.261/H.263、MPEG-1、MPEG-2、MPEG-4和H.264等。第三部分讲述各种数字电视标准，包括ATSC、OpenCable™、DVB和ISDB以及IPTV的相关概念。

本书内容系统完整，概念叙述清楚严谨，可作为视频通信、数字电视、视频服务等领域广大研发和应用人员的参考手册，也可供高等院校有关专业师生阅读。

图灵电子与电气工程丛书

视频技术手册（第5版）

◆ 著　［美］Keith Jack
译　杨　征　田尊华　张杰良　陈　虎　等
责任编辑　朱　巍

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街14号
邮编　100061　电子函件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京隆昌伟业印刷有限公司印刷

◆ 开本：787×1092　1/16
印张：42.5
字数：1170千字　2009年8月第1版
印数：1-3 000册　2009年8月北京第1次印刷

著作权合同登记号　图字：01-2008-3323号

ISBN 978-7-115-21020-3/TN

定价：119.00元

读者服务热线：（010）51095186　印装质量热线：（010）67129223

反盗版热线：（010）67171154

版权声明

Video Demystified, Fifth Edition by Keith Jack, ISBN: 978-0-7506-8395-1.

Authorized Simplified Chinese translation edition published by the Proprietor.
ISBN: 978-981-272-111-2.

Elsevier (Singapore) Pte Ltd.
3 Killiney Road
#08-01 Winsland House I
Singapore 239519
Tel：(65)6349-0200
Fax：(65)6733-1817

First Published 2009
2009年初版



译　者　序

对于视频，我们并不陌生，日常生活中经常看的电影、电视就是视频的典型应用。人类接收的信息中约有70%来自视觉，周围景物在人眼视网膜上的映像是人类最有效、最重要的信息交流方式。视频具有确切、直观、具体生动、真实和高效的特点，这正是视频技术的无穷魅力所在。

随着计算机软硬件技术的迅猛发展，视频技术也得到了飞速发展。现在，数字娱乐、多媒体通信、高清晰度电视、数字电视、宽带网络等都是围绕如何有效地交流视频信息而展开研究的。但是，由于视频信息十分丰富且信息量大，故对视频信号的处理、传输、存储和显示等都提出了新的要求。因此，视频技术的研究和应用是目前信息技术领域最热门的话题之一。

10多年之前，Keith Jack撰写的*Video Demystified*一书回答了全球范围内的工程师提出的各种视频相关的问题，在业内广受好评。时至今日，人民邮电出版社购买了该书第5版的版权，并委托我们翻译，我们很高兴地接受了这一任务。

与前几版相比，本书的内容更加充实，不仅增加了对最新视音频编解码器的介绍，而且更新了所有广播、有线和卫星规范，以及从模拟到数字电视信号转换的有关信息。在视音频编码器部分中，增加了对MPEG-4.2、MPEG-4.10（H.264）和SMPTE 421M（VC-1）的介绍；在消费类电子设备中，增加了对HD DVD、蓝光播放器和录像机、数字媒体适配器、数字机顶盒、数字电视、网络电视（IPTV）机顶盒、便携式媒体播放器和移动视频接收器的介绍，并讲述了支持这些设备的相关技术，如隐藏字幕、字幕、图文电视与V-Chip技术、先进的视频处理技术、复杂的图像合成技术、基于IP网络的ARIB与DVB技术和数字版权管理（DRM）。与前几版一样，本书条理清晰，深入浅出，紧密联系当前视频技术领域已取得的最新成果和未来发展方向，系统地讲述视频技术的基础理论和实践技能，这对读者了解和掌握视频技术有很大的帮助。

本书的翻译工作由国防科技大学从事计算机、多媒体和通信等相关专业工作、具有丰富视频技术经验的研究人员承担，主要由杨征（第1、10、12、13、14章）、田尊华（第2章到第7章、第20章）、张杰良（第8、9和11章）和陈虎（第15到第19章）翻译，杨征对全书进行了统稿。此外，参与本书翻译的人员还有：闫志强、岳虹、张波涛、刘齐军、林龙信、李晋文、肖枫涛、张聪、韩智文、马蓉、焦贤龙、邝祝芳、奚丹、刘志忠、陈钢、宋锐、陈威兵、唐玲艳、唐扬斌、叶俊、杨明军、颜炯、薄建禄、肖国尊、何小威、杨磊、冯权友、伍微、齐宁。Be Flying工作室（http://blog.csdn.net/be_flying）负责人肖国尊对本书的翻译和出版做了大量的协调和规范工作，在此予以衷心感谢。

视频技术是一门交叉性的学科，发展十分迅速。由于译者的水平和学识有限，译本中不尽妥当之处在所难免，恳请读者批评指正。

目　录

第1章 引　　言

几年前，视频的应用还略微局限于模拟方面，主要用于广播和有线电视、VCR、机顶盒、电视和便携式摄像机等。之后，模拟视频大量而迅速地转向数字视频，这种转变主要基于MPEG-2视频压缩标准。

目前，除了传统的DV、MPEG-1和MPEG2音频和视频压缩标准外，还出现了三种高性能的视频压缩标准。在相同的视频质量下，这些新的编解码器能提供更高的压缩比。

- ❑ MPEG-4.2。这种视频编解码器在压缩比方面比MPEG-2提高了1.5~2倍。它本来能够占领更多的市场份额，但是由于其复杂性，实际上一直没有得到广泛的认可。而且，很多人也决定干脆等待新的MPEG-4.10（H.264）编解码器的问世。
- ❑ MPEG-4.10（H.264）。这种编解码器在压缩比方面比MPEG-2提高了2~3倍。而且，随着编码器的逐步完善且支持MEPG-4.10（H.264）的工具越来越多，MEPG-4.10（H.264）的压缩比和质量也将不断提升。吸取了MPEG-4的教训，MPEG-4.10（H.264）已经能够适应于低价位、单芯片的产品，并且被DVB标准和ARIB标准所采纳。
- ❑ SMPTE 421M（VC-1）。这是MPEG-4.10（H.264）的有力竞争者。这种视频编解码器在压缩比方面同样比MPEG-2提高了2~3倍，而且，随着编码器的逐步完善，压缩比和压缩质量也将进一步得到改善。

随着人们对音频相关技术的兴趣不断增加，越来越多的音频编解码技术也随之出现。这些音频相关技术包括：6.1和7.1声道音频技术、多声道无损压缩技术、相同音频质量情况下的低比特率技术以及在需要较高音质且比特率比较合理的条件下为某些应用实现最高音频质量的技术。

除了音频解码外，DVD、HD DVD、蓝光录像机和数字视频录像机（Digital Video Recorder，DVR）也需要使用实时、高质量的音频编码技术。综合这些对音频技术的需求可见，消费市场上的单芯片解决方案通常需要集成DSP用于音频处理。

消费者使用的电子设备也变得更成熟了，能支持更多类型的内容和内部互连方式。目前，主要有以下相关的消费类电子设备。

- HD DVD、蓝光播放器和录像机。除了用来播放CD和DVD，这些先进的HD播放器同样支持播放MPEG-4.10（H.264）和SMPTE 421M（VC-1）格式的内容。有些还包含以太网连接功能，从而能够在电视机上非常便捷地欣赏个人电脑或媒体服务器中的内容。
- 数字媒体适配器。这些小型、低价的盒子使用以太网或802.11连接功能，从而使个人电脑或媒体服务器上的内容能够很容易地在电视机上播放。同时，它们通常支持MPEG-2、MPEG-4.10（H.264）、SMPTE 421M（VC-1）以及JPEG格式的内容。
- 数字机顶盒。有线、卫星机顶盒现在已具有数字视频录像机（DVR）的功能，使得浏览者能很方便地随时观看他们想看的内容。MPEG-4.10（H.264）和SMPTE 421M（VC-1）的使用，有利于增加内容频道，减少数字产品的过早过时。
- 数字电视（Digital Television，DTV）。电视中除集成了调谐器和解码器，甚至在一些电视中还包含数字媒体适配器功能。此外，支持在线观看视频的技术也在不断发展。
- 网络电视（IPTV）机顶盒。这些低价位的机顶盒在那些拥有高速DSL和光纤到户（FTTH，

Fiber To The Home）功能的地区很受欢迎。MPEG-4.10（H.264）和SMPTE 421M（VC-1）的使用降低了数字产品过早过时的可能性。

- *便携式媒体播放器*。通过使用内置硬盘驱动器（Hard Disc Drive，HDD），这些播放器利用USB接口或802.11网络接口实现与电脑的连接，用于下载多种类型的内容。便携式媒体播放器通常也支持MPEG-2、MPEG-4.10（H.264）、SMPTE 421M（VC-1）以及JPEG格式内容的播放。
- *移动视频接收器*。这种设备通常集成在手机内，其中MPEG-4.10（H.264）和SMPTE 421M（VC-1）用来传送高质量的视频信号。例如，DMB、DVB-H和DVB-SH标准就属于这类应用。

当然，为了生产这类先进的消费产品，不仅需要音频、视频编解码技术的支持，还需要如下技术的支持。

- *隐藏字幕、字幕、图文电视与V-Chip技术*。为了支持数字广播，这些标准已做了更新。
- *先进的视频处理技术*。由于视频内容以及视频显示器的分辨率多种多样，导致通常需要高质量的缩放与运动自适应的逐行扫描技术。因为标清（SD）电视标准和高清（HD）电视标准使用的是不同的色度标准，所以在HDTV上观看标清内容或者在SDTV上观看高清内容都需要进行修正。
- *复杂的图像合成技术*。该技术能够用来渲染复杂的图像，在该图像中可能包含多种视频、屏幕视控系统（On-Screen Display，OSD）、字幕/子图、文本和图表等元素。
- *基于IP网络的ARIB与DVB技术*。随着基于IP网络的ARIB与DVB技术逐步应用，支持IP视频的复杂性也在提高。
- *数字版权管理*（Digital Rights Management，DRM）。该技术可以用来防止内容被未授权复制和未授权观看。

本书进行了一些更新，以反映上述这些技术的发展。实现一个实际的视频解决方案并不容易，而且许多工程师对该领域知之甚少或者没有足够的实践经验。对于那些需要理解视频特性并在新一代设计中实现视频技术的工程师，本书无疑是他们了解和使用视频技术的指南。

本书同样适用于需要或渴望了解视频技术的工程师、正在研发新视频产品的VLSI设计工程师，以及任何想要评估或了解视频系统的人士。

1.1 本书内容

本书内容按以下方式组织。

第2章是视频简介，讨论了各种视频格式和信号及其应用领域，描述了隔行扫描视频与逐行扫描视频之间的区别。这一章还给出了DVD播放器和数字机顶盒的结构框图。

第3章回顾了常见的颜色空间，说明了各种颜色空间在数学上的关系，以及何时使用某种特定的颜色空间。这一章所讨论的颜色空间包括RGB、YUV、YIQ、YCbCr、xvYCC、HSI、HSV和HLS。此外，这一章还讨论了如何从一个非RGB颜色空间变换到RGB颜色空间，以及如何进行gamma修正。

第4章是视频信号概述，回顾了视频时序以及各种视频格式的模拟信号和数字信号表示方法，这些视频格式包括480i、480p、576i、576p、720p、1080i和1080p。

第5章讨论了消费类应用和专业视频应用中的模拟视频接口，包括模拟RGB、YPbPr、S-Video和SCART接口。

第6章讨论了半导体、专业视频设备和消费类设备中的各种数字视频接口，回顾了BT.601和

BT.656半导体接口，SDI、SDTI和HD-SDTI专业视频接口，以及DVI、HDMI和IEEE 1394消费产品类视频接口。

第7章讨论了几种数字视频处理技术，例如从YCbCr 4：4：4到YCbCr 4：2：2变换、YCbCr数字滤波器模板、视频缩放、隔行/逐行转换、帧率转换、alpha混合、闪烁过滤和色度抠像。此外，这一章还介绍了亮度、对比度、饱和度、色调和锐度控制等。 3

第8章概述了NTSC、PAL和SECAM。这一章首先回顾了各种复合模拟视频信号格式和视频测试信号。这一章所讨论的VBI数据包括时间码、隐藏字幕和扩展数据服务（Extended Data Service，XDS）、宽屏信令与图文电视。此外，这一章还回顾了PALplus、RF调制、BTSC、Zweiton模拟立体声以及NICAM 728数字立体声技术。

第9章讨论了应用于NTSC和PAL制式彩色视频信号编解码的数字技术。同时，这一章还回顾了亮度/色度（Y/C）分离技术并进行了比较分析。

第10章讨论了用于远程视频会议的H.261和H.263视频压缩标准。

第11章讨论了便携式数字摄像机中使用的消费电子类DV视频压缩标准。

第12章回顾了MPEG-1视频压缩标准。

第13章讨论了MPEG-2视频压缩标准。

第14章讨论了MPEG-4视频压缩标准，其中包括MPEG-4.10（H.264）。

第15章讨论了美国使用的ATSC数字电视标准。

第16章讨论了美国使用的OpenCable™数字电视标准。

第17章讨论了欧洲和亚洲使用的DVB数字电视标准。

第18章讨论了日本使用的ISDB数字电视标准。

第19章讨论了IPTV。该技术可以实现通过Internet、DSL、FTTH等广播网络来传送压缩视频。

最后，第20章给出了400多个视频方面的术语。如果遇到不熟悉的术语，可以在这个术语表中找到其定义。 4

1.2 标准组织

许多标准化组织参与了视频标准的制定，下面列出了部分组织。

美国先进电视系统委员会（ATSC）：www.atsc.org

日本无线工业及商贸联合会（ARIB）：www.arib.or.jp

美国有线电视实验室：www.cablelabs.com

美国家电行业团体消费电子协会（CEA）：www.ce.org

数字视频广播（DVB）：www.dvb.org

美国电子工业协会（EIA）：www.eia.org

欧洲广播联盟（EBU）：www.ebu.ch

欧洲电信标准协会（ETSI）：www.etsi.org

国际电工委员会（IEC）：www.iec.ch

美国电气电子工程师学会（IEEE）：www.ieee.org

国际标准化组织（ISO）：www.iso.org

国际电信联盟（ITU）：www.itu.int

电缆电信工程师协会（SCTE）：www.scte.org

运动图像和电视工程师协会（SMPTE）：www.smpte.org

视频电子标准协会（VESA）：www.vesa.org 5

第2章 视频简介

尽管视频信号有很多类型和实现技术，但其目的只是为了实现视觉信息在不同位置之间的传送。这些视觉信息可能是来自于VCR、DVD播放器、本地广播的某个频道、有线电视，或者来自于卫星电视系统、因特网，还可能是其他各种渠道。

但有一点是肯定的，那就是视频信息必须从一个设备传输到另一个设备。可能是从卫星电视机顶盒或DVD播放器传输到电视机，或者在卫星电视机顶盒或电视机内部不同芯片之间传输。这看起来好像很简单，但需要满足很多不同的要求，故存在很多方式来实现这种传输。

2.1 模拟与数字

几年前，大多数的视频设备还主要是为模拟视频而设计的，数字视频还仅限于视频编辑这样的专业应用领域。

目前，普通消费者在日常生活中都用上了数字视频设备，这要得益于这些数字产品价格的不断下降。这种趋势也导致相关技术的快速发展，如DVD播放机和摄像机、数字机顶盒、数字电视（DTV）、便携视频播放器和基于因特网的视频数据传输能力等。

2.2 视频数据

起初，视频仅仅包含灰度（也称为黑白）信息。

在建立彩色广播电视系统的过程中，人们试图用模拟RGB（红、绿、蓝）来发送彩色视频。然而，这种技术占用的带宽是当时使用的灰度解决方案的3倍多，因此必须创建其他的替代方法。于是，人们用Y、R-Y和G-Y数据来表示颜色信息，并开发相应的技术来传输Y、R-Y和G-Y信息。这种技术只需要一个信号，而不是3个独立的信号，因此只需要与原来传输灰度视频信号相同的带宽。今天广泛使用的NTSC、PAL和SECAM视频标准仍然是建立在这种复合视频信号（composite video signal）基础上的。在第8章和第9章还将更加详细地对其进行讨论。

现在，尽管有很多种表示视频的方式，但它们都要通过RGB的数学公式表示出来。在第3章将更加详细地讨论这些表示方法。

S-Video是为了将终端设备连接（目的不是为了广播）在一起而建立的。每组信号由两个模拟信号构成：一个为灰度（Y）信息，另一个是以特定格式传输的模拟R-Y和B-Y颜色信息（也称为C或色度）。这种技术曾经只在S-VHS中可用，但现在大多数消费类视频产品都支持这项技术。在第9章将更加详细地对其进行讨论。

尽管模拟RGB视频数据一直用于专业视频市场，但为了实现与高端消费设备的连接，模拟RGB视频数据也暂时应用于普通视频市场。与S-Video一样，模拟RGB视频数据也不用于广播。

Y、R-Y、G-Y视频信号的另一种形式称为YPbPr，现在通常用于将消费类视频产品连接在一起。其主要优点在于能够在消费类视频产品之间传输高分辨率视频。有些制造商错误地将YPbPr连接器称为YUV、YCbCr或Y（B-Y）（R-Y）。

第5章将详细地讨论各种模拟视频互连方案。

2.2.1 数字视频

目前最通用的数字信号为RGB和YCbCr。RGB是模拟RGB视频信号进行简单数字化后得到的版本。YCbCr基本上是模拟YPbPr视频信号的数字化版本，这种格式由DVD和数字电视所采用。

第6章将更加深入地讨论各种数字视频互连方案。

2.2.2 最佳连接方法

设备的最佳连接方法是什么？对于DVD播放器和数字有线电视/卫星电视/地面机顶盒，按照视频质量由好到差排序，一般的顺序为：

(1) HDMI（数字YCbCr）
(2) HDMI（数字RGB）
(3) 模拟YPbPr
(4) 模拟RGB
(5) 模拟S-Video
(6) 模拟复合视频

有些人可能不同意这个排序。但是，大多数消费类产品都是在YCbCr颜色空间进行数字视频处理的。因此，使用YCbCr作为设备互连格式可以减少所需的颜色空间转换次数。数字信号的颜色空间转换仍然优先选择进行D/A（数字到模拟）转换，接着进行A/D（模拟到数字）转换，所以HDMI RGB排在模拟YPbPr之前。

计算机产业已经对接入计算机显示器的模拟和数字RGB信号进行了标准化。

2.3 视频时序

视频画面尽管看起来好像是连续运动的，其实那是一系列静止的图像。这些图像切换得足够快，使得画面看起来像是连续运动的，如图2-1所示。对于消费类电子产品，通常情况下为每秒50或60张图像，计算机显示器为每秒70~90张图像。一种称为场同步（vertical sync）的特定时序信息被用于指定新图像从什么时候开始显示。

每张静止图像是由扫描线（scan line）组成的，即沿着显示器从上到下、一行接着一行进行显示的数据线，如图2-1所示。另一种称为行同步（horizontal sync）的时序信息用于指定新扫描线什么时候开始显示。

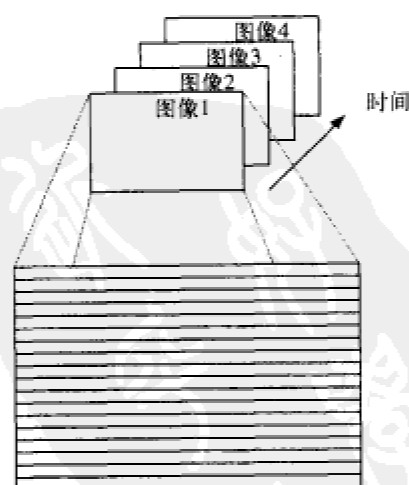


图2-1 视频是由一系列静止图像组成的，每张图像是由多个独立的数据线组成的

7

行同步和场同步信息通常通过以下三种方式之一进行传输：

(1) 单独的行同步和场同步信号；
(2) 单独的复合同步信号；
(3) 嵌入视频信号的复合同步信号。

复合同步信号是由场同步和行同步信号组合而成的。

使用模拟RGB视频的计算机和消费类设备通常采用技术（1）或（2）。支持复合视频或模拟

YPbPr视频的消费类设备通常采用技术（3）。

对于数字视频，通常要么采用技术（1），要么就是将时序编码字嵌入数字视频流中。这个问题将在第6章详细讨论。

隔行与逐行

由于视频是由一系列静止图像组成的，因此，简单地连续显示每幅完整的图像是有意义的，也就是一张接一张地显示每幅图像。

上面所讲的就是逐行显示（progressive）［或称为非隔行（non-interlaced）显示］的基本技术。对于将图像以逐行方式“绘制”到屏幕上的设备（如CRT），每张图像都是从显示器的左上角开始，一直向右移动，直到到达显示器的右边缘为止，然后向下扫描一行，重复地从左到右进行扫描。这个过程一直持续到整个屏幕全部被刷新一次为止，如图2-2所示。

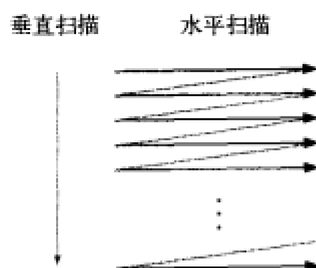


图2-2 逐行显示器一行接一行地连续“绘制”出每幅图像

8

在电视发展的早期，采用一种称为“隔行显示（interlacing）”的技术来减少每幅图像所需发送的数据量。这种方式先发送奇数行的数据，接着发送偶数行的数据（如图2-3所示），因此每次发送的数据量为一幅图像数据量的一半。

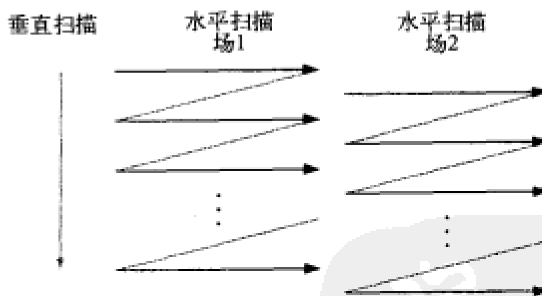


图2-3 隔行显示器先“绘制”图像的一半（奇数行），然后“绘制”图像的另一半（偶数行）

既然隔行显示方式有这个优点，那么为什么还要多余地使用逐行显示方式呢？

在隔行显示方式中，每个扫描线更新的频率只有同等情况下逐行显示方式的一半。所以，为了避免由于过低的帧率而导致颜色对比显著的边缘出现闪烁，要限制扫描线之间的变化，这本质上是由于在垂直方向上对图像进行了低通滤波。而逐行显示方式对线之间的变化就没有这种限制，因此，逐行显示方式能够提供更高分辨率（在垂直方向上）的没有闪烁的图像。

现在，大多数广播（包括HDTV）都仍然以隔行方式发射信号。大多数基于CRT的显示器也是隔行方式显示，而LCD、等离子以及计算机显示器都是采用逐行方式。

2.4 视频分辨率

在现实生活中，人们对于视频分辨率的概念往往有一些模糊的认识。我们通常见到的视频分辨率为720×480或1920×1080。然而，那只是水平采样和垂直扫描线的数目，并不等于说必须具有这么多有用信息。

例如，可以以13.5MHz的频率对模拟视频信号进行采样，生成每行720个采样点。对同样的信号以27MHz的频率采样可以生成每行1440个采样点。但这仅仅是每行的采样数目不同，视频内容的分辨率并没有改变。

因此，视频质量通常用*线分辨率*（lines of resolution）来度量。本质上是表示在显示器上可以显示多少不同的黑白垂直线。然后将这个数目归一化为1：1显示宽高比（对于4：3的显示器，将该数除以3/4，而对于16：9的显示器则除以9/16）。当然，对于宽屏（16：9）显示器，这会使得视频分辨率较低，与直觉不相符。

2.4.1 标准清晰度

标准清晰度（standard-definition）视频通常定义为480或576的隔行有效扫描线的视频，分别称为“480i”和“576i”。

固定像素（非CRT）、具有4：3宽高比的消费类显示器通过转换后，对应于720×480i或720×576i的有效分辨率。而对于16：9的宽高比，转换后的有效分辨率对应于960×480i或960×576i。

2.4.2 增强清晰度

增强清晰度（enhanced-definition）视频通常定义为具有480或576逐行有效扫描线的视频，分别称为“480p”或“576p”。

固定像素（非CRT）具有4：3宽高比的消费类显示器转换后的有效分辨率对应于720×480p或720×576p。而对于16：9的宽高比，转换后的有效分辨率对应于960×480p或960×576p。

9

标准清晰度和增强清晰度的差异在于，标准清晰度是隔行的，而增强清晰度是逐行的。

2.4.3 高清晰度

高清晰度（high-definition）视频通常定义为具有720逐行（720p）或1080隔行（1080i）有效扫描线的视频。固定像素（非CRT）、具有16：9宽高比的消费类显示器转换后的有效分辨率分别为1280×720p或1280×1080i。

但是，HDTV显示器在技术上定义为能够显示最少720p或1080i有效扫描线。它还必须能够用至少540逐行（540p）或810隔行（810i）有效扫描线来显示16：9的视频画面。这样，在制造具有4：3宽高比、基于CRT的HDTV和具有16：9宽高比、分辨率为1024×1024p、1280×768p、1024×768p的LCD/等离子等显示器时，能够降低制造费用。

2.5 音频与视频压缩

近年来，数字电视、DVD播放机和摄像机、数字视频摄像机等消费类电子产品取得了一些最新进展，这归功于音频和视频压缩标准。这些压缩标准主要有：带有Dolby® Digital的MEPG-2、DTS、MPEG-1或MPEG-2音频。

新的音频和视频编解码器，像MPEG-4 HE-AAC、MPEG-4.10（H.264）和SMPTE 421M（VC-1），在保持同样质量时，提供了比以往编解码方法更高的压缩效率。这些进展使得可以使用新的视频发布方式（面向消费者和家庭内部）、新的消费产品（如便携视频播放器和移动视频/移动电话）和更多的有线/卫星电视频道。

2.6 应用框图

下面几个简化的框图有助于我们认识视频流是如何经过各种相应处理的。

2.6.1 DVD播放器

图2-4为基本DVD播放机的简化框图，其中显示的是公共模块。如今所有这些模块都集成在一个廉价的芯片上了。

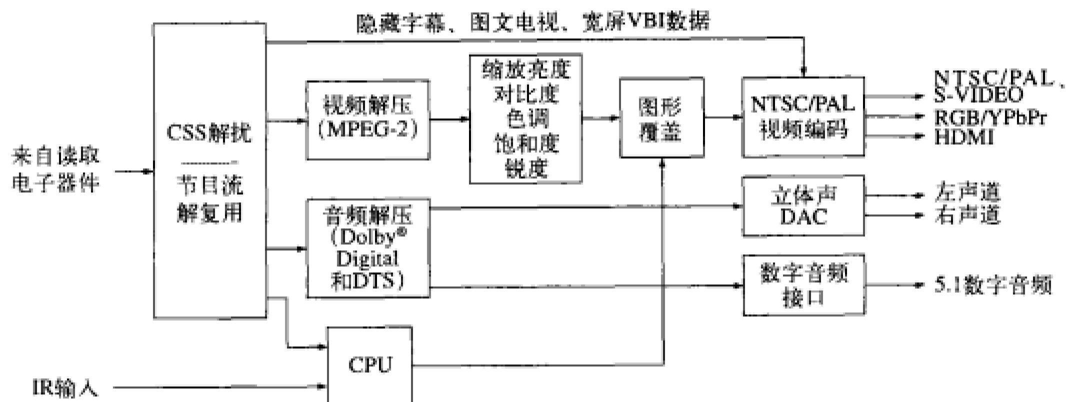


图2-4 简化的基本DVD播放机框图

除了播放DVD（基于MPEG-2视频压缩技术）外，目前，DVD播放器还有望处理MP3和WMA音频、MPEG-4视频（DivX视频）、JPEG图像等。同时有望具有特定的播放模式，如以各种速度慢/快进和快/慢退等。对DVD音频和SACD的支持也是很流行的。

DVD播放机的一个新增功能是能够接入到家庭网络中，用于播放一些存放于PC中的媒体（音乐、视频、图像等）。这些“网络化的DVD播放机”也包括播放一些准备事后浏览的内容，如来源于因特网的电影和下载到内部硬盘（HDD）中的媒体。支持对各种闪存卡中的音频、视频和图像进行播放的这类需求也在不断增长。

有些DVD播放机制造商为了快速吸引买家的注意力，对视频的频率响应进行了超频处理，使得其产品看起来与众不同。但因为这种特性通常需要长时间的激励过程，所以很容易失败或需要调试。对于观看电影来讲，很多视频爱好者都认为频率响应应该尽可能平滑。

10~11

另一个问题是模拟视频信号的量化等级。尽管很容易产生每个精确的视频量化等级，但这些等级变化得相当快。目前有些评论也指出了这个问题，因为在信号源之间切换时可能会使发生变化的亮度或灰度值具有平均作用，从而使得用户进行的任何校正或手动调整失效。

2.6.2 数字媒体适配器

数字媒体适配器是连接到家庭网络、用于播放存放于PC或媒体存储器中的多媒体内容（音乐、视频、图像等）的设备。这些小而廉价的盒子使得媒体内容很容易在家庭中的任意或所有电视机上欣赏。许多数字媒体适配器支持无线网络功能，这就简化了数字媒体适配器的安装过程。

图2-5是基本的数字媒体适配器的简化框图，图中显示的是公共模块。如今所有这些模块都集成在一个廉价的芯片上了。

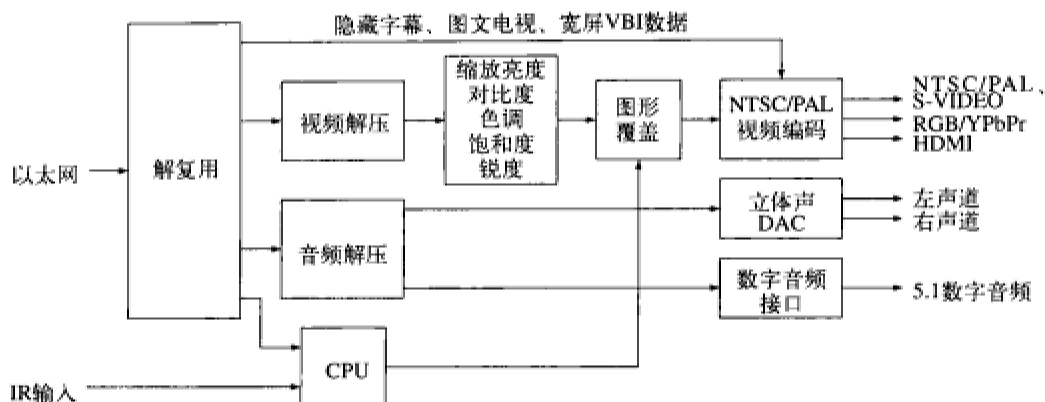


图2-5 简化的数字媒体适配器框图

2.6.3 数字电视机顶盒

数字电视标准可以分为7大类：

(1) ATSC（美国先进电视系统委员会）

(2) DVB（数字视频广播）

(3) ARIB（日本无线工业及商贸协会）

(4) IPTV（网络电视，包括基于IP的DVB和ARIB）

(5) 开放数字有线标准，如OpenCable标准

(6) 带版权保护的数字有线标准

(7) 带版权保护的数字卫星电视标准 12

起初这些机顶盒都是基于MPEG-2视频和Dolby®Digital或MPEG音频，目前的机顶盒支持新的先进音视频标准，如MPEG-4 HE-AAC音频、Dolby®Digital Plus音频、MPEG-4.10（H.264）视频和SMPTE（VC-1）视频。

图2-6是数字电视机顶盒的简化结构图，图中显示的是公共音视频处理模块。数字机顶盒用于接收数字电视广播，这些数字电视广播可能来自地面台站（空中电波）、有线或卫星。在数字电视内部就可能包含这些电路。

目前的许多电视机顶盒有两个调谐器，并具有数字视频摄像（VCR）功能。这使得在浏览一个内部HDD时，还可以往另一个内部HDD中读入节目。在数字电视接收机中，有了这两个调谐器，还可以支持PIP特性（即将2个输入影像输出到同一画面上）。 13

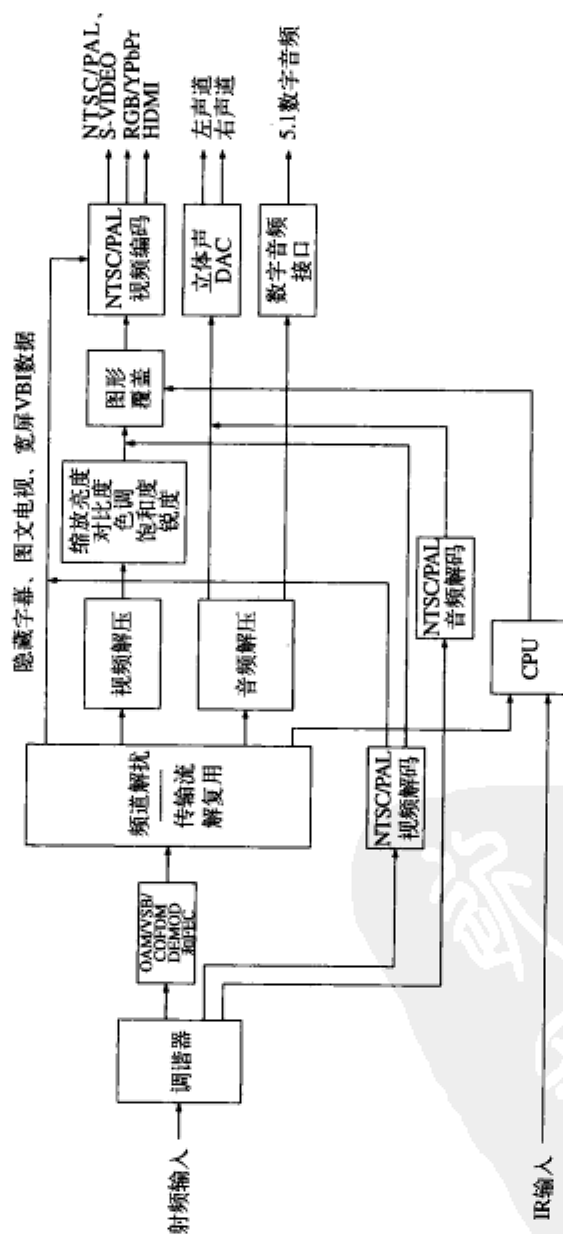


图2-6　数字电视机顶盒的简化框图

第3章 颜色空间

颜色空间（color space）是颜色集合的数学表示。三个最常用的颜色模型是：RGB（用于计算机图形学中）；YIQ、YUV或YCbCr（用于视频系统中）；CMYK（用于彩色打印）。但是，这些颜色空间都与色度、饱和度和亮度这些直观概念没有直接的关系。这使得人们暂时去寻求其他的模型，如HSI和HSV，来简化编程、处理和终端用户操作。

所有的颜色空间都能够由RGB信息推导而来，像摄像机和扫描仪这样的设备都采用RGB信息。

3.1 RGB颜色空间

红、绿和蓝（RGB）颜色空间广泛应用于计算机图形学和显示器。红绿蓝是三种基本的加性颜色（各个颜色成分相加来形成所需的颜色），可以用三维的笛卡尔坐标系统来表示RGB颜色空间（见图3-1）。图中立方体的对角线所代表颜色的每个基本颜色分量都相等，对角线上不同的点具有不同的灰度值。表3-1包含的RGB值具有100%的幅度、100%饱和彩条（color bar）和公共的视频测试信号。

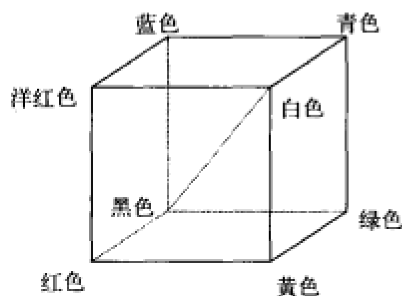


图3-1 RGB颜色立方体

表3-1 100%RGB彩条

	标称范围	白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
R	0~255	255	255	0	0	255	255	0	0
G	0~255	255	255	255	255	0	0	0	0
B	0~255	255	0	255	0	255	0	255	0

RGB颜色空间是计算机图形学最通常的选择，因为彩色显示器采用红、绿和蓝三色来生成目标颜色。因此，选择RGB颜色空间简化了系统的架构与设计。同时，由于RGB颜色空间已经使用了很多年，因此，采用RGB颜色空间设计的系统可以利用大量已有的软件模块。

但是，当处理现实世界图像时，使用RGB颜色空间并不是很有效。要在RGB颜色立方体中生成任何一种颜色，三个RGB分量都需要占用相同的带宽。这就使得每个RGB颜色分量的帧缓冲需要同样的像素深度和显示分辨率。同时，在RGB颜色空间处理图像通常也不是最有效的方法。例如，为了修改给定像素的亮度或颜色值，必须同时从帧缓冲中读出RGB颜色的三个分量，然后重新计算亮度和颜色，执行相应的修改，再计算新的RGB值并写回帧缓冲。如果能够访问到直接以亮度和颜色格式存储的图像，那么某些处理步骤会更快。

由于这些和其他的某些原因，许多视频标准采用亮度（luma）和两个色差信号。最常用的就是YUV、YIQ和YCbCr颜色空间。尽管所有这些颜色空间都是相关的，但是它们之间还是有一些差异的。

3.1.1 sRGB

RGB颜色空间存在许多不同类型的实现，sRGB颜色空间是对这些实现的规范化。与sRGB（IEC61966-2-1）相关的规范使用BT.709色度、D65参考白色、2.2显示色彩增益（gamma）和线性RGB（每个颜色8位）。

sRGB值具有归一化的颜色范围，即颜色值处于0到1之间。8位数字sRGB表示的黑白色值范围为0~255。其中有一个称为Studio RGB的版本使用8位表示16~235范围内的黑白色，这就可以与视频应用兼容。

sRGB有一个不足之处，由于归一化值限定于0~1范围内，在这个范围（生成的颜色三角形）之外的颜色就不可用。为此建立了扩展RGB颜色空间scRGB。

16

3.1.2 scRGB

scRGB颜色空间（更正式的称呼是sRGB64）扩展了sRGB的动态范围、颜色范围和位精度。scRGB的范围不仅比sRGB要大得多，而且比人类视觉系统可以看到的颜色范围还要大。scRGB（IEC 61966-2-2）规定：使用BT.709色度、D65参考白色和线性RGB数据（每个颜色16位）。

除了使用归一化颜色范围0~1外，同时也支持颜色范围－0.5~＋7.4999。尽管scRGB与sRGB具有相同的基本颜色，但由于可以使用小于0和大于1的颜色值，scRGB具有更大的颜色表示范围。线性16位scRGB值与归一化颜色范围之间的关联关系为：

00000＝－0.5
04096＝0.0（黑色）
12288＝1.0（白色）
16384＝1.5
65535＝7.4999

在进行gamma校准之后，非线性的16位scR′G′B′值与归一化颜色范围的关联关系为：

00000＝－0.7354
04096＝0.0（黑色）
12288＝1.0（白色）
65535＝2.3876

3.1.3 scRGB颜色到sRGB颜色的转换



将线性16位scRGB颜色转换为经过gamma校准的8位sRGB颜色（表示为$sR'G'B'_8$）：

$scR = (scR_{16}/8192) - 0.5$
$scG = (scG_{16}/8192) - 0.5$
$scB = (scB_{16}/8192) - 0.5$

如果 $(scR_{16}, scG_{16}, scB_{16}) \leqslant 4095$

$sR'_8 = 0$
$sG'_8 = 0$
$sB'_8 = 0$

如果 $4096 \leqslant (scR_{16}, scG_{16}, scB_{16}) \leqslant 4243$

$sR'_8 = round[4.500 \times scR \times 255]$
$sG'_8 = round[4.500 \times scG \times 255]$
$sB'_8 = round[4.500 \times scB \times 255]$

如果 $4244 \leqslant (scR_{16}, scG_{16}, scB_{16}) \leqslant 12288$

$sR'_8 = \text{round}[(1.099 \times scR^{0.45} - 0.099) \times 255]$

$sG'_8 = \text{round}[(1.099 \times scG^{0.45} - 0.099) \times 255]$

$sB'_8 = \text{round}[(1.099 \times scB^{0.45} - 0.099) \times 255]$

如果（$scR_{16}, scG_{16}, scB_{16}$）$\geqslant 12289$

$sR'_8 = 255$

$sG'_8 = 255$

$sB'_8 = 255$

3.2 YUV颜色空间

YUV颜色空间是下面三大复合颜色视频标准所采用的颜色空间：PAL（逐行倒相彩色电视制式）、NTSC（美国国家电视系统委员会制式）和SECAM（按序传送彩色与存储制式）。黑白系统仅使用亮度（Y）信息，颜色信息（U和V）以一种特定的方式加入，使得黑白电视接收机同样可以显示标准的黑白图像。而彩色电视接收机则对额外的彩色信息进行解码从而显示彩色画面。

gamma校准的RGB（表示为R′G′B′）与YUV颜色空间之间的转换公式为：

17

$$
\begin{aligned}
Y &= 0.299R' + 0.587G' + 0.114B' \\
U &= -0.147R' - 0.289G' + 0.436B' \\
&= 0.492(B' - Y) \\
V &= 0.615R' - 0.515G' - 0.100B' \\
&= 0.877(R' - Y) \\
R' &= Y + 1.140V \\
G' &= Y - 0.395U - 0.581V \\
B' &= Y + 2.032U
\end{aligned}
$$

对于取值范围在0~255的数字R′G′B′值，Y的取值范围为0~255，U的取值范围为0~±122，V的取值范围为0~±157。通常可以对这些转换公式进行裁减以简化在实际的NTSC或PAL数字编码器和解码器中的实现。

注意，对于数字数据，8位YUV和R′G′B′数据应该在0和255处饱和，以避免向下溢出和向上溢出导致的反转问题。

如果使用了（B′－Y）和（R′－Y）的整个取值范围，那么复合NTSC和PAL的量化等级将超出（当前的）黑白电视发送器和接收器所能支持的取值范围。实验表明，调制的次波在白色以上和黑色以下的偏移达到亮度信号（Y）偏移的20%是准许的。那么需要选择缩放因子使得最大的量化等级对应于75%的幅度、100%的饱和黄色和青色彩条，即这个颜色为最亮的白色（100 IRE）。

3.3 YIQ颜色空间

YIQ颜色空间（将在第8章深入讨论）是由YUV颜色空间推导而来的，NTSC复合颜色视频标准有选择性地使用YUV颜色空间。（I代表“同相”，Q代表“正交”，它们指的是用于发射颜色信息的调制方法。）在R′G′B′和YIQ颜色之间转换的基本公式为：

$$
\begin{aligned}
Y &= 0.299R' + 0.587G' + 0.114B' \\
I &= 0.596R' - 0.257G' - 0.321B' \\
&= V\cos 33^\circ - U\sin 33^\circ \\
&= 0.736(R' - Y) - 0.268(B' - Y) \\
Q &= 0.212R' - 0.523G' + 0.311B' \\
&= V\sin 33^\circ - U\cos 33^\circ \\
&= 0.478(R' - Y) + 0.413(B' - Y)
\end{aligned}
$$

$R' = Y + 0.956I + 0.621Q$

$G' = Y - 0.272I + 0.647Q$

$B' = Y - 1.107I + 1.704Q$

对于取值范围在0~255的数字R'G'B'值，Y的取值范围为0~255，I的取值范围为0~±152，Q的取值范围为0~±134。I和Q是通过将U和V轴旋转33°而获得的。这些公式通常可以进行裁减，来简化实际NTSC数字编码器和解码器的实现。

18 注意，对于数字数据，8位的YIQ和R'G'B'数据应该在量化等级0和255处达到饱和，以避免向下和向上溢出而导致反转问题。

3.4　YCbCr颜色空间

随着世界范围内数字分量视频标准（将在第4章讨论）的发展，YCbCr颜色空间作为ITU-R BT.601的一部分发展起来了。YCbCr颜色空间是YUV颜色空间的缩放和偏移版本。Y定义为具有8位，标称颜色表示范围为16~235；Cb和Cr标称颜色表示范围定义为16~240。我们将对几种YCbCr采样格式进行描述，如4：4：4、4：2：2、4：1：1和4：2：0。

3.4.1　RGB与YCbCr颜色空间之间的转换公式：SDTV

1. RGB到YCbCr颜色空间的转换：模拟公式

许多规范假定源信号为模拟R'G'B'，并具有0~1的归一化颜色范围。这样，首先要转换为模拟YCbCr颜色空间：

$Y = 0.299R' + 0.587G' + 0.114B'$

$Pb = -0.169R' - 0.331G' + 0.500B'$

$Pr = 0.500R' - 0.419G' - 0.081B'$

为了生成8位正确的YCbCr颜色，需要将模拟YCbCr颜色量化为8位的颜色值：

$Y = \mathrm{round}[219Y + 16]$

$Cb = \mathrm{round}[224Pb + 128]$

19 $Cr = \mathrm{round}[224Pr + 128]$

2. RGB到YCbCr颜色空间的转换：数字公式

为了将标称范围16~235（Studio R'G'B'）内的8位数字R'G'B'颜色转换为YCbCr颜色，必须对模拟公式进行简化：

$Y = 0.299R' + 0.587G' + 0.114B'$

$Cb = -0.172R' - 0.339G' + 0.511B' + 128$

$Cr = 0.511R' - 0.428G' - 0.083B' + 128$

3. YCbCr到RGB颜色空间的转换：模拟公式

很多规范假定信号源为模拟YPbPr颜色空间。这首先得转换为模拟R'G'B'颜色：

$R' = Y + 1.402Pr$

$G' = Y - 0.714Pr - 0.344Pb$

$B' = Y + 1.772Pb$

为了生成标称取值范围为16~235（Studio R'G'B'）的8位R'G'B'颜色，将R'G'B'颜色量化为8位：

$out' = \mathrm{round}[219in' + 16]$

4. YCbCr到RGB颜色空间的转换：数字公式

为了将标称取值范围为16~235（Studio R'G'B'）的8位YCbCr数据转换为R'G'B'颜色，模拟公式可以简化为：

$R' = Y + 1.371\ (Cr - 128)$
$G' = Y - 0.689\ (Cr - 128) - 0.336\ (Cb - 128)$
$B' = Y + 1.732\ (Cb - 128)$

5. YCbCr到RGB颜色空间转换的一般注意事项

当进行YCbCr到RGB的颜色空间转换时，得到的R′G′B′颜色值的标称取值范围为16~235，取值可能偶尔会偏移到0~15和236~255的范围内，这是由于视频处理和噪声的缘故，Y和CbCr分别会偶尔超出16~235和16~240的范围。注意，8位YCbCr和R′G′B′数据应该在0和255颜色量化等级处达到饱和，以避免向下和向上溢出导致反转问题。

表3-2列出了75%幅度、100%饱和彩条（这是一个通用的视频测试信号）对应的YCbCr颜色值。

表3-2 75%YCbCr彩条

	标称范围	白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
SDTV									
Y	16~235	180	162	131	112	84	65	35	16
Cb	16~240	128	44	156	72	184	100	212	128
Cr	16~240	128	142	44	58	198	212	114	128
HDTV									
Y	16~235	180	168	145	133	63	51	28	16
Cb	16~240	128	44	147	63	193	109	212	128
Cr	16~240	128	136	44	52	204	212	120	128

6. 关于计算机系统的注意事项

如果R′G′B′数据的取值范围为0~255（在计算机系统中通常是这样的），下面的公式可能更便于使用：

$Y = 0.257R' + 0.504G' + 0.098B' + 16$
$Cb = -0.148R' - 0.291G' + 0.439B' + 128$
$Cr = 0.439R' - 0.368G' - 0.071B' + 128$
$R' = 1.164(Y - 16) + 1.596(Cr - 128)$
$G' = 1.164(Y - 16) - 0.813(Cr - 128) - 0.391(Cb - 128)$
$B' = 1.164(Y - 16) + 2.018(Cb - 128)$

注意，8位的YCbCr和R′G′B′数据在0和255颜色量化等级处应该达到饱和，以避免由于向下和向上溢出导致反转问题。

3.4.2 RGB与YCbCr颜色空间之间的转换公式：HDTV

1. RGB到YCbCr颜色空间转换：模拟公式

许多规范都假定源信息为标称范围为0~1的模拟R′G′B′。首先要转换为模拟YPbPr颜色：

$Y = 0.213R' + 0.715G' + 0.072B'$
$Pb = -0.115R' - 0.385G' + 0.500B'$
$Pr = 0.500R' - 0.454G' - 0.046B'$

为了生成8位正确的YCbCr颜色值，需要将YCbCr颜色量化为8位：

$Y = \mathrm{round}[219Y + 16]$
$Cb = \mathrm{round}[224Pb + 128]$
$Cr = \mathrm{round}[224Pr + 128]$

2. RGB到YCbCr颜色空间转换：数字公式

为了将标称取值范围为16~235（Studio R′G′B′）的8位数字R′G′B′数据转换为YCbCr颜色数据，可以将模拟公式进行简化：

$Y = 0.213R' + 0.715G' + 0.072B'$

$Cb = -0.117R' - 0.394G' + 0.511B' + 128$

$Cr = 0.511R' - 0.464G' - 0.047B' + 128$

3. YCbCr到RGB颜色空间转换：模拟公式

许多规范假定源信息为模拟YPbPr。首先要转换为模拟R′G′B′颜色值：

$R' = Y + 1.575Pr$

$G' = Y - 0.468Pr - 0.187Pb$

$B' = Y + 1.856Pb$

为了生成标称取值范围为16~235的8位R′G′B′颜色值（Studio R′G′B′），将R′G′B′量化为8位数据：

$out' = round[219in' + 16]$

4. YCbCr到RGB颜色空间转换：数字公式

为了将8位YCbCr数据转换为标称值为16~235的R′G′B′颜色值（Studio R′G′B′），对模拟公式进行简化：

$R' = Y + 1.540(Cr - 128)$

$G' = Y - 0.459(Cr - 128) - 0.183(Cb - 128)$

$B' = Y + 1.816(Cb - 128)$

5. YCbCr到RGB颜色空间转换的一般注意事项

当进行YCbCr到R′G′B′颜色空间的转换时，得到的R′G′B′颜色值的标称取值范围为16~235，可能偶尔会偏移到0~15和236~255的范围内。这是因为，由于视频处理和噪声的原因，Y和CbCr的取值分别偶尔会超出16~235和16~240的范围。注意，8位YCbCr到R′G′B′数据在颜色量化等级0和255处应该达到饱和，以避免向下和向下溢出而导致反转问题。

表3-2列出了75%幅度、100%饱和彩条（这是一个通用的视频测试信号）对应的YCbCr颜色值。

6. 关于计算机系统的注意事项

如果R′G′B′数据的取值范围为0~255（在计算机系统中通常是这样的），下面的公式可能更便于使用：

$Y = 0.183R' + 0.614G' + 0.062B' + 16$

$Cb = -0.101R' - 0.338G' + 0.439B' + 128$

$Cr = 0.439R' - 0.399G' - 0.040B' + 128$

$R' = 1.164(Y - 16) + 1.793(Cr - 128)$

$G' = 1.164(Y - 16) - 0.534(Cr - 128) - 0.213(Cb - 128)$

$B' = 1.164(Y - 16) + 2.115(Cb - 128)$

注意，8位的YCbCr和R′G′B′数据在0和255颜色量化等级处应该达到饱和，以避免由于向下和向上溢出导致反转问题。

3.4.3　4：4：4 YCbCr格式

图3-2表示的是4：4：4格式YCbCr采样点的定位。每个采样点有Y、Cb和Cr值，每个颜色值的各颜色分量具有典型的8位（消费类应用）或10位（专业视频应用）。因此每个采样点需要24位

（对于专业视频应用则为30位）。

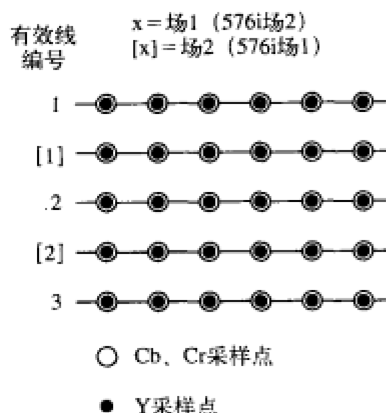


图3-2 4：4：4协调位置采样。采样位置位于隔行图像的有效扫描线上

3.4.4 4：2：2 YCbCr格式

图3-3表示的是4：2：2格式YCbCr采样点的定位。对于每两个水平Y采样点，有一个Cb和一个Cr采样点。每个采样颜色值的各颜色分量具有典型的8位（消费类应用）或10位（专业视频应用）。因此每个采样点需要16位（对于专业视频应用则为20位），通常采用图3-4的格式。

为了显示4：2：2的YCbCr数据，首先要转换为4：4：4的YCbCr数据，这需要采用插值来生成丢失的Cb和Cr采样信息。

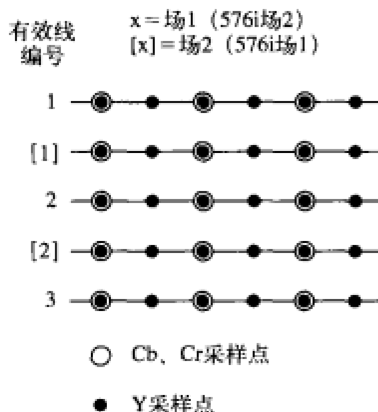


图3-3 4：2：2协调位置采样。采样位置位于隔行图像的有效扫描线上

采样点 0	采样点 1	采样点 2	采样点 3	采样点 4	采样点 5
Y7-0	Y7-1	Y7-2	Y7-3	Y7-4	Y7-5
Y6-0	Y6-1	Y6-2	Y6-3	Y6-4	Y6-5
Y5-0	Y5-1	Y5-2	Y5-3	Y5-4	Y5-5
Y4-0	Y4-1	Y4-2	Y4-3	Y4-4	Y4-5
Y3-0	Y3-1	Y3-2	Y3-3	Y3-4	Y3-5
Y2-0	Y2-1	Y2-2	Y2-3	Y2-4	Y2-5
Y1-0	Y1-1	Y1-2	Y1-3	Y1-4	Y1-5
Y0-0	Y0-1	Y0-2	Y0-3	Y0-4	Y0-5
CB7-0	CR7-0	CB7-2	CR7-2	CB7-4	CR7-4
CB6-0	CR6-0	CB6-2	CR6-2	CB6-4	CR6-4
CB5-0	CR5-0	CB5-2	CR5-2	CB5-4	CR5-4
CB4-0	CR4-0	CB4-2	CR4-2	CB4-4	CR4-4
CB3-0	CR3-0	CB3-2	CR3-2	CB3-4	CR3-4
CB2-0	CR2-0	CB2-2	CR2-2	CB2-4	CR2-4
CB1-0	CR1-0	CB1-2	CR1-2	CB1-4	CR1-4
CB0-0	CR0-0	CB0-2	CR0-2	CB0-4	CR0-4

每个采样点16位

-0＝采样点0数据
-1＝采样点1数据
-2＝采样点2数据
-3＝采样点3数据
-4＝采样点4数据

图3-4 4：2：2帧缓冲区的格式

3.4.5 4：1：1 YCbCr格式

图3-5表示的是4：1：1格式的YCbCr（也就是我们熟知的YUV12格式）采样点的定位。这种

格式主要应用于某些消费类视频和DV（数字摄像机）视频压缩应用中。对于每4个水平Y采样点，有一个Cb和一个Cr采样点。每个采样点的颜色分量具有典型的8位数据。因此每个采样点需要12位，通常采用图3-6的格式。

为了显示4：1：1的YCbCr数据，首先要转换为4：4：4的YCbCr数据，这需要采用插值来生成丢失的Cb和Cr采样信息。

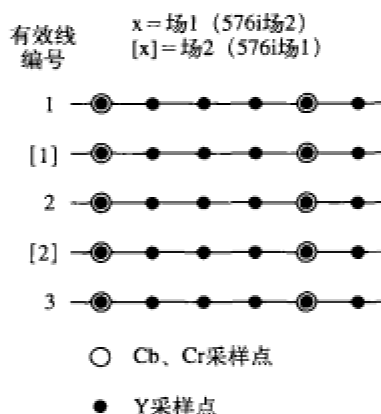


图3-5　4：1：1协调位置（co-sited）采样。采样位置位于隔行图像的有效扫描线上



图3-6　4：1：1帧缓冲区的格式

3.4.6　4：2：0 YCbCr格式

与4：2：2 YCbCr仅在水平方向对Cb和Cr减少为2：1不同，4：2：0 YCbCr在水平和垂直方向都减少为2：1。这主要是应用于视频压缩领域。

22 ~ 23

如图3-7~图3-11所示，有多种4：2：0 YCbCr格式。表3-3列举了各种DV应用中的YCbCr数据格式。

表3-3　针对各种DV应用的YCbCr格式

YCbCr格式	25Mbit/s的DV					50Mbit/s的DV			100Mbit/s的DV		MPEG-1	MPEG-2、-4.2、-4.10（H.264）	H.21、H.23
	480线的DV	560线的DV	480线的DVCAM	576线的DVCAM	D-7(DVCPRO)	DVCPRO 50	数字Betacam	D-9（数字S）	DVCPRO HD	D-9 HD			
4：4：4 协调位置												×	
4：2：2 协调位置						×	×	×	×	×		×	
4：1：1 协调位置	×		×		×								
4：2：0											×	×	×
4：2：0 协调位置		×		×									

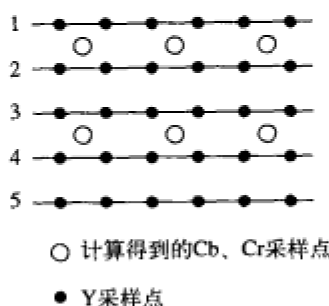


图3-7 针对H.261、H.263和MPEG-1的4：2：0采样。采样点位于逐行或非隔行图像的有效扫描线上

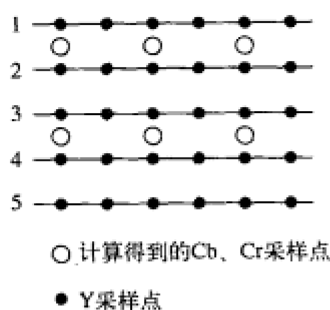


图3-8 针对MPEG-2、MPEG-4.2、MPEG-4.10（H.264）和SMPTE 421M（VC-1）的4：2：0采样。采样点位于逐行或非隔行图像的有效扫描线上

24

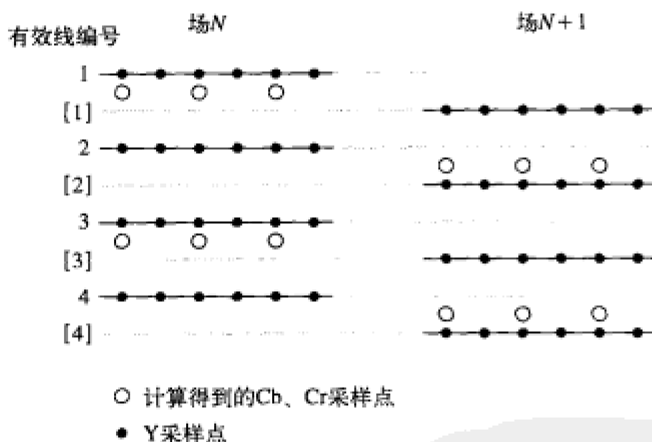


图3-9 针对MPEG-2、MPEG-4.2、MPEG-4.10（H.264）和SMPTE 421M（VC-1）的4：2：0采样。采样点位于隔行图像的有效扫描线上（图中上面的域从1开始，top_field_first = 1）

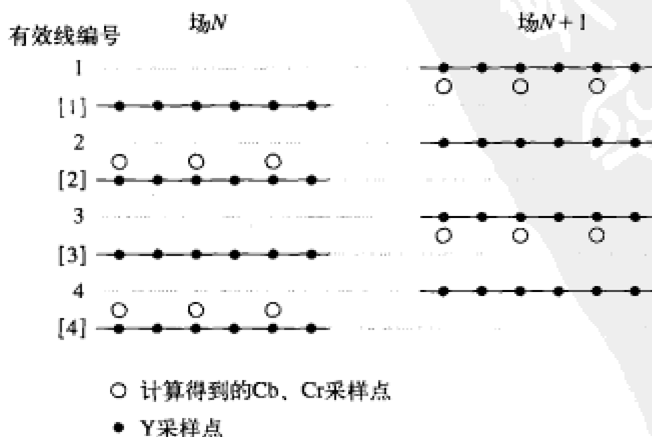


图3-10 针对MPEG-2、MPEG-4.2、MPEG-4.10（H.264）和SMPTE 421M（VC-1）的4：2：0采样。采样点位于隔行图像的有效扫描线上（图中上面的域从0开始，top_field_first = 0）

25

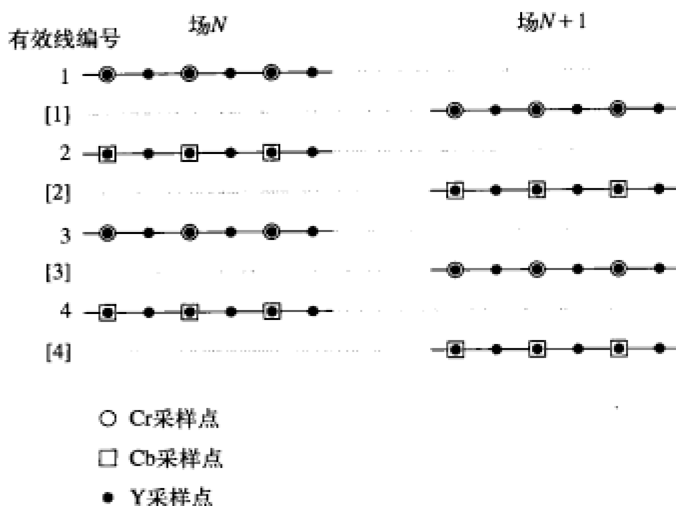


图3-11　针对576i DV和DVCAM的4：2：0协调位置采样。采样点位于隔行图像的有效扫描线上

为了显示4：2：0的YCbCr数据，首先要转换为4：4：4的YCbCr数据，这需要采用插值来生成丢失的Cb和Cr采样信息。注意，有些解并不能正确地将4：2：0的YCbCr数据转换为4：4：4格式的YCbCr数据，因而会导致“色度陷阱”。

3.5　xvYCC颜色空间

xvYCC（用于视频的扩展色域YCbCr）颜色空间是对基本YCbCr的色域进行扩展后得到的颜色空间，它使得颜色数目多了1.8倍。xvYCC规范（IEC61966-2-4）使用BT.709色度标准和D65参考白色。scR′G′B′和xvYCC颜色空间之间的转换公式与R′G′B′和YCbCr之间的转换公式是一样的。

基于xvYCC的YCbCr颜色数据有8位，表示范围为1~254，这使得可以与现有的设计向后兼容。Y有8位，表示范围为－15/219~238/219（－0.068493~1.086758）；CbCr有8位，表示范围为－15/224~238/224（－0.066964~1.062500）。HDMI用色域边界描述（gamut boundary description）元数据传输正在使用的xvYCC视频数据。

3.6　PhotoYCC颜色空间

PhotoYCC（美国柯达公司的一个商标）是为了对Photo CD图像数据进行编码而开发的颜色标准。目的就是要建立一个与显示设备无关的颜色空间。为了最大化视频显示效率，PhotoYCC颜色空间是建立在ITU-R BT.601和BT.709基础上的。

编码过程（RGB到Photo YCC颜色空间的转换）假定使用CIE标准光源（CIE Standard Illuminant）D_{65}，并且图像采集系统的光谱灵敏度与BT.709参考原色系的颜色匹配函数成比例。与计算机图形系统的RGB颜色值不同，在此RGB颜色值可能为负数。PhotoYCC还包含有BT.709色域之外的颜色；这些额外的颜色值使用负值进行编码。

26

3.6.1　RGB到PhotoYCC的颜色空间转换

线性RGB数据（归一化到0~1范围内）到PhotoYCC的非线性变换方法如下：

对于 (R,G,B)≥0.018

$R' = 1.099R^{0.45} - 0.099$

$G' = 1.099G^{0.45} - 0.099$

$B' = 1.099B^{0.45} - 0.099$

对于 $-0.018 < (R,G,B) < 0.018$

$R' = 4.5R$

$G' = 4.5G$

$B' = 4.5B$

对于 $(R,G,B) \leq -0.018$

$R' = -1.099|R|^{0.45} - 0.099$

$G' = -1.099|G|^{0.45} - 0.099$

$B' = -1.099|B|^{0.45} - 0.099$

由取值范围为0~255的R′G′B′颜色值，可以生成一个亮度信号和两个色度信号（C1和C2）：

$Y = 0.213R' + 0.419G' + 0.081B'$

$C1 = -0.131R' - 0.256G' + 0.387B' + 156$

$C2 = 0.373R' - 0.312G' - 0.061B' + 137$

作为一个例子，20%的灰度值（R、G和B都等于0.2）在PhotoCD磁盘中记录的值为：

$Y = 79$

$C1 = 156$

$C2 = 137$

3.6.2 PhotoYCC到RGB的颜色空间转换

由于PhotoYCC试图保留电影的动态范围，对PhotoYCC图像进行编码需要选择适合输出设备的颜色空间和取值范围。因此，解码公式并不总是编码公式的准确逆过程。下面的公式适合于生成驱动CRT显示器的RGB颜色值，并假定在编码图像和显示图像中的亮度值具有统一的关系：

$R' = 0.981Y + 1.315(C2 - 137)$

$G' = 0.981Y - 0.311(C1 - 156) - 0.669(C2 - 137)$

$B' = 0.981Y + 1.601(C1 - 156)$

R′G′B′颜色值应该在0~255范围之内，否则就达到饱和。上面的公式假定，显示器使用的荧光色度（phosphor chromaticity）与BT.709参考原色系相同，并且视频信号亮度（V）和显示亮度（L）满足如下关系：

对于 $V \geq 0.0812$

$L = ((V + 0.099)/1.099)^{1/0.45}$

对于 $V < 0.0812$

$L = V/4.5$

3.7 HSI、HLS和HSV的颜色空间转换

HSI（色调、饱和度、亮度）和HSV（色调、饱和度、纯度）颜色空间在处理颜色时更加“直观”，在设计上使得颜色感知和解释的方式与人很接近。在需要手动指定颜色值时经常需要使用它们，但是，当用户可以以可视化的方式选择颜色或指定潘通专业色卡（Pantone color）时，就很少使用它们了。主要是从传统习惯角度来讨论这些颜色空间。HLS（色调、亮度和饱和度）类似于HSI，只是使用了术语lightness，而不是intensity（光强度）来表示亮度。

HSI和HSV颜色空间之间的差异在于亮度分量（I或V）的计算方式，亮度分量同时确定了亮度（I或V）和饱和度（S）的分布和动态取值范围。HSI颜色空间最适合于传统的图像处理函数，如卷积、均化、直方图等，可以通过处理亮度值来实现这些操作，因为亮度I对R、G、B值的依赖程度是一样的。HSV颜色空间适合于处理色度和饱和度（颜色偏移或调整颜色数量），因为这

使得饱和度具有更大的动态取值范围。

图3-12表示的是一个单六角锥HSV颜色模型。六角锥的上面对应于V＝1，即具有最大亮度值的颜色。六角锥下面的顶点为黑色，V＝0。互补色以H为测度，围绕垂直轴（V）的角度相差180°，即相互处于六角锥对等的位置上，红色对应的角度为0°。S的取值为一个比率（ratio），取值范围从0到1，0对应于垂直轴（V）的中心线，1对应于六角锥的边。S在0到1之间的任何取值都与V＝0的点相关。S＝0、V＝1的点为白色。S＝0，V的中间值为灰色。注意，当S＝0时，H的值是不相关的。从艺术家的观点来看，任何V＝1、S＝1的颜色都是纯颜料（其颜色由H来定义）。增加白色等于减小S（V保持不变）；增加黑色等于减小V（S保持不变）。色调（tone）是通过同时减小S和V来生成的。表3-4列出了幅度为75%、饱和度为100%的HSV彩条。

表3-4　75%HSV彩条

	标称取值范围	白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
H	0°~360°	–	60°	180°	120°	300°	0°	240°	–
S	0~1	0	1	1	1	1	1	1	0
V	0~1	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0

图3-13表示的是双六角锥HSI颜色模型。六角锥的上面对应于I＝1，即白色。六角锥下面的顶点为黑色，对应于I＝0。互补色以H为测度，围绕垂直轴（I）的角度相差180°，即相互处于六角锥对等的位置上，红色对应的角度为0°（为了与HSV模型一致，我们已经改变了泰克公司（Tektronix）在0°处为0的约定）。S的取值范围从0到1，0对应于垂直轴（I），1对应于六角锥的面。灰度S恒等于0，但色度的最大饱和值在S＝1、I＝0.5处。表3-5列出了幅度为75%、饱和度为100%的HSI彩条。

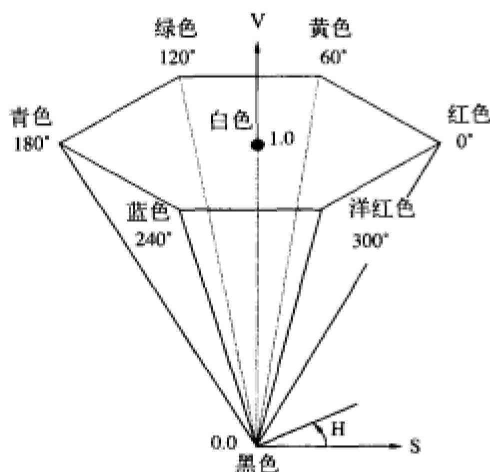


图3-12　单六角锥HSV颜色模型

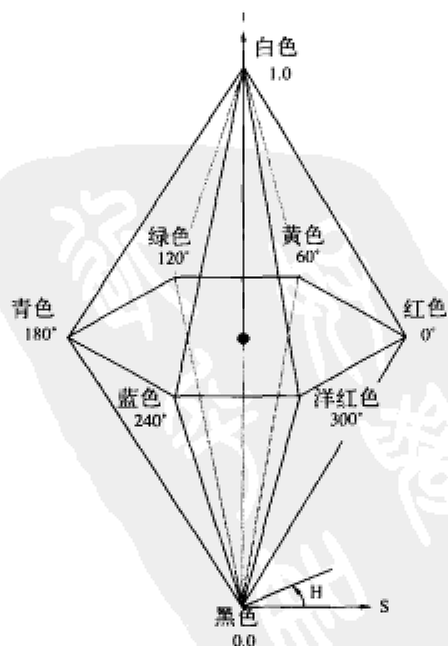


图3-13　双六角锥HSI颜色模型。为了与HSV模型兼容，我们已经改变了泰克公司的蓝色在0°处的约定，将该模型描述为双六角锥，而不是双圆锥

表3-5 75%HSV彩条。为了与HSV模型兼容，我们已经改变了泰克公司的蓝色在0°处的约定

	标称取值范围	白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
H	0°~360°	–	60°	180°	120°	300°	0°	240°	–
S	0~1	0	1	1	1	1	1	1	0
I	0~1	0.75	0.375	0.375	0.375	0.375	0.375	0.375	0

3.8 色度图

图3-14给出了人类正常的视觉（1931 CIE标准观察者）能感受到的色域。该图及其潜在的数学表达式分别在1960年和1976年更新过；然而，NTSC电视系统是基于1931规范的。

颜色感知是通过观察由三个CIE（国际照明委员会）基本颜色的组合来度量的，这三个基本颜色为：波长为700nm的红色、波长为546.1nm的绿色和波长为435.8nm的蓝色。这三个基本颜色和其他谱系是由这三个基本色混合而来的纯色，在图3-14中是沿着曲线外边界线（称为光谱轨迹）上的颜色。

光谱轨迹的端点（位于红色和蓝色处）是由一条直线连接起来的，这条直线代表紫色，紫色是由红色和蓝色构成的组合颜色。在这个封闭边界内的区域包含了所有可以通过混合不同颜色的光而生成的颜色。距离边界越近的颜色，其饱和度越大。处于边界线内的颜色越接近图的中心（白色），感觉越柔和。图中的每一个点代表唯一的一种颜色，可以用其x和y坐标来唯一指定。

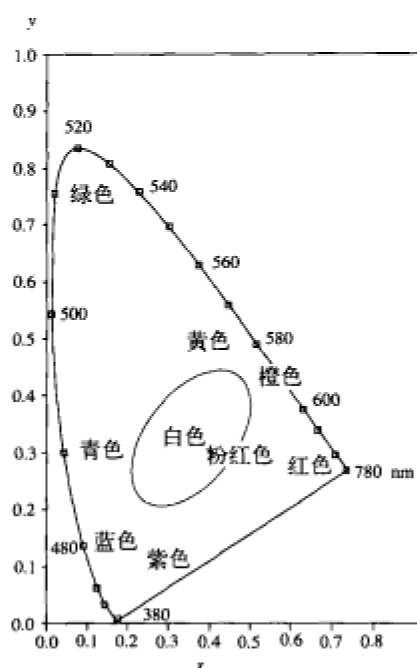


图3-14 CIE 1931色度图，表示各种颜色区间

28 ≀ 30

在CIE系统内，红、绿和蓝的亮度被转换成被称为*三色值*（tristimulus value）的量，并用大写字母X、Y和Z表示。这些值表示基本色的相对量。

图3-14中的坐标轴是由三色值推导而来：

$$x = X/(X+Y+Z)$$
$$= \text{red}/(\text{red}+\text{green}+\text{blue})$$
$$y = Y/(X+Y+Z)$$
$$= \text{green}/(\text{red}+\text{green}+\text{blue})$$
$$z = z/(X+Y+Z)$$
$$= \text{blue}/(\text{red}+\text{green}+\text{blue})$$

坐标x、y和z称为*色度坐标*，三者之和恒等于1。因此，z总可以用x和y来表示，为了表示任何颜色，只需要指定x和y的值，相应的图也可以是二维的。

31

典型情况下，信号源或显示器指定三个（x, y）坐标来定义要使用的三个基本色。由这三个（x, y）坐标形成的三角形围成的色域为该信号源或显示器所能产生的颜色。如图3-15所示，图中对NTSC、PAL和HDTV的色域进行了比较。注意，这三种颜色组成的集合并不能生成所有可能的颜色，这也是为什么电视图像的颜色总是不完全精确的。

另外，信号源或显示器通常要使用白色的坐标（*x*，*y*），因为纯白色通常并不能采集到或再生。白色定义为采集到或产生的、所有三个基本色都相等的、并具有细微颜色阴影的颜色。注意，亮度（luminance或brightness）信息并不包含在标准的CIE 1931色度图中，而是垂直于（*x*，*y*）平面的一个轴。颜色越亮，色度取值范围越小。

1953 NTSC标准的RGB色度和参考白色（CIE亮度C）为：

R：$x_r=0.67$　$y_r=0.33$
G：$x_g=0.21$　$y_g=0.71$
B：$x_b=0.14$　$y_b=0.08$
白色：$x_w=0.3101$　$y_w=0.3162$

现代的NTSC、480i和480p视频系统使用的是不同的RGB色度集合（SMPTE "C"）和参考白色（CIE亮度D_{65}）：

R：$x_r=0.630$　$y_r=0.340$
G：$x_g=0.310$　$y_g=0.595$
B：$x_b=0.155$　$y_b=0.070$
白色：$x_w=0.3127$　$y_w=0.3290$

PAL、SECAM、576i和576p视频系统使用的RGB色度和参考白色（CIE亮度D_{65}）为：

R：$x_r=0.64$　$y_r=0.33$
G：$x_g=0.29$　$y_g=0.60$
B：$x_b=0.15$　$y_b=0.06$
白色：$x_w=0.3127$　$y_w=0.3290$

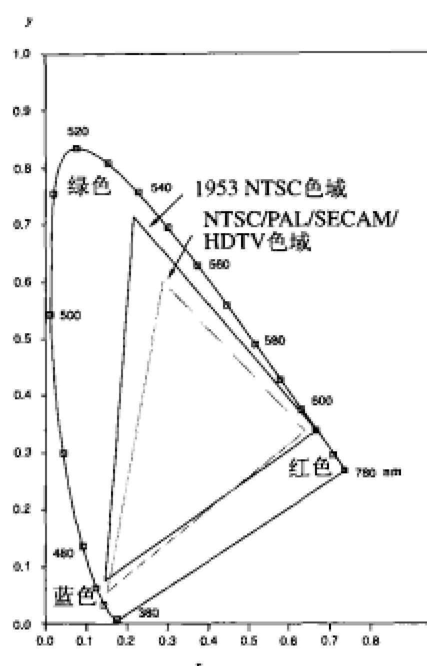


图3-15　CIE 1931色度图，显示了各种色域

sRGB、scRGB、xvYCC和HDTV视频系统使用的RGB色度和参考白色（CIE亮度D_{65}）是基于BT.709和SMPTE 274M的：

R：$x_r=0.64$　$y_r=0.33$
G：$x_g=0.30$　$y_g=0.60$
B：$x_b=0.15$　$y_b=0.06$
白色：$x_w=0.3127$　$y_w=0.3290$

由于各种视频标准使用不同的色度和参考白色值，当信号源和显示器的值不匹配时会发生细微的颜色误差；例如，在HDTV上播放480i或480p节目，或在NTSC电视机上播放HDTV节目。通过使用一个3×3的矩阵乘法器，这些细微的颜色误差很容易在显示器上得以校正，我们将在第7章讨论这些内容。

消费类显示器的RGB色度通常与这些标准稍微有所不同。导致的结果是，一种或多种RGB颜色有细微的偏移，如在红色中有过多的橙色，或者在绿色中有过多的蓝色。这通常可以通过对显示器进行专业校准得以补偿。

3.9　关于非RGB颜色空间的注意事项

在处理非RGB颜色空间（如YIQ、YUV或YCbCr）的信息时，必须注意，有些值并不能组合在一起，否则就可能产生无效的RGB颜色值。术语无效（invalid）是指RGB颜色分量超出了归一化的RGB颜色范围（1,1,1）。

例如，给定RGB的归一化值（1,1,1），得到的YCbCr值为（235,128,128）。如果对Cb和Cr值进行处理来产生一个YCbCr值（235,64,64），那么对应的RGB归一化值为（0.6,1.29,0.56）。注意，绿色值超出了归一化值1。

由这个例子可以很明显地看出，Y、Cb和Cr的很多组合都可能导致无效的RGB值，这些YCbCr值必须经过处理以生成有效的RGB颜色值。图3-16表示的是将RGB归一化界限转换为YCbCr颜色空间。

图3-16 RGB界限转换为三维YCbCr颜色空间

使用恒定亮度和色度的方法可以得到最好的结果：当Cb和Cr限制在最大有效值（这些有效值与到达界限之前的无效颜色具有同样的色度）处时，Y不变。恒定色度原则就是将无效的CbCr组合移到CbCr的原点（128,128），直到它们位于有效的YCbCr颜色块表面为止。

当从非RGB颜色空间转换到RGB颜色空间时，必须注意，要考虑到饱和处理逻辑以确保由于数字电路的有限精度而导致向上和向下溢出的反转条件不会发生。8位RGB颜色值小于0时一定要设为0，超过255的颜色值要设为255。

33

3.10 gamma校准

大多数CRT显示器的变换函数产生的亮度值正比于信号幅度的某种能量（称为gamma）。因此，高亮度的范围被扩展了，而低亮度的范围被压缩了（见图3-17）。在抑制噪声方面，这是一个优点，这是因为人眼对相等的相对亮度变化具有差不多相同的敏感性。在发射之前对视频信号进行gamma校准，显示器的亮度输出就大体上是线性的了（图3-17中的灰线），并且发射过程中产生的噪声也会受到抑制。

为了最大程度地降低图像较暗区域的噪声，现代的视频系统对黑色区域的曲线增益做了限制。这种技术对靠近黑色的增益做了限制，而对曲线的其他部分进行扩展，以保持函数和正切的连续性。

尽管视频标准假定显示器的gamma大约为2.2，但对于CRT显示器来讲，gamma值取2.5更加现实。然而，这种差异改善了光线较暗环境下的视觉效果。在明亮光照环境下更准确的视觉效果是通过采用另一个gamma因子来完成的，这个gamma因子的取值大约为1.14（2.5/2.2）。为了接近“电影观赏效果（film look）”，在显示器中通常对gamma曲线进行微调（tweak）。

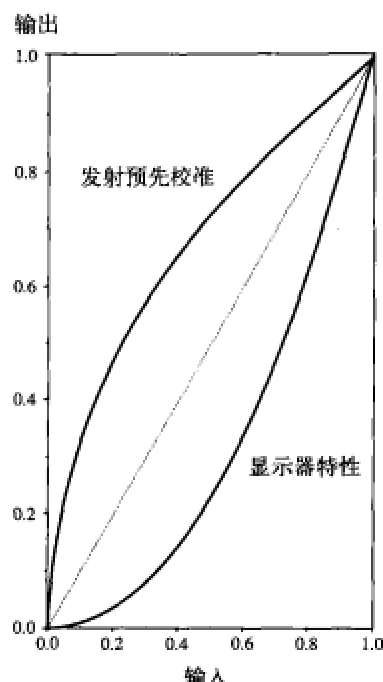


图3-17 gamma校准的效果

3.10.1 早期的NTSC系统

早期的NTSC系统假定在显示器中会有一个简单的变换，其gamma值为2.2。RGB值归一化为取值范围0~1：

$R = R'^{2.2}$

$G = G'^{2.2}$

$B = B'^{2.2}$

为了对非线性显示器进行补偿，线性RGB数据在发射之前要经过逆变换的方式进行“gamma校准”。RGB值要归一化为取值范围0~1：

$R' = R^{1/2.2}$

$G' = G^{1/2.2}$

$B' = B^{1/2.2}$

34

3.10.2 早期的PAL和SECAM系统

大多数早期的PAL和SECAM系统都假定在显示器中有一个简单的变换，变换的gamma值为2.8。RGB值归一化为取值范围0~1：

$R = R'^{2.8}$

$G = G'^{2.8}$

$B = B'^{2.8}$

为了对非线性显示器进行补偿，线性RGB数据在发射之前要经过逆变换的方式进行“gamma校准”。RGB值要归一化为取值范围0~1：

$R' = R^{1/2.8}$

$G' = G^{1/2.8}$

$B' = B^{1/2.8}$

3.10.3 当前的系统

当前的NTSC、480i、480p和HDTV视频系统假定在显示器中进行下面的变换，变换的gamma值为[1/0.45]。RGB值归一化为取值范围0~1：

如果 $(R',G',B') < 0.081$

$R = R'/4.5$

$G = G'/4.5$

$B = B'/4.5$

如果 $(R',G',B') \geq 0.081$

$R = ((R' + 0.099)/1.099)^{1/0.45}$

$G = ((G' + 0.099) / 1.099)^{1/0.45}$

$B = ((B' + 0.099) / 1.099)^{1/0.45}$

扩展域颜色空间（如scRGB）对小于0的值要进行额外的处理：

如果 $(R', G', B') < -0.081$

$R = -((R' - 0.099) / -1.099)^{1/0.45}$

$G = -((G' - 0.099) / -1.099)^{1/0.45}$

$B = -((B' - 0.099) / -1.099)^{1/0.45}$

如果 $-0.081 \leqslant (R', G', B') < 0.081$

$R = R'/4.5$

$G = G'/4.5$

$B = B'/4.5$

为了对非线性显示器进行补偿，线性RGB数据在发射之前要经过逆变换的方式进行“gamma校准”。RGB值要归一化为取值范围0~1：

如果 $(R, G, B) < 0.081$

$R' = 4.5R$

$G' = 4.5G$

$B' = 4.5B$

如果 $(R, G, B) \geqslant 0.081$

$R' = 1.099\ R^{0.45} - 0.099$

$G' = 1.099\ G^{0.45} - 0.099$

$B' = 1.099\ B^{0.45} - 0.099$

扩展域颜色空间（如scRGB）对小于0的值需要进行额外的处理：

如果 $(R, G, B) < -0.081$

$R' = -1.099\ (-R^{0.45}) + 0.099$

$G' = -1.099\ (-G^{0.45}) + 0.099$

$B' = -1.099\ (-B^{0.45}) + 0.099$

如果 $-0.081 \leqslant (R', G', B') < 0.081$

$R' = 4.5R$

$G' = 4.5G$

$B' = 4.5B$

35

尽管大多数PAL和SECAM标准规定gamma取2.8，但现在通常采用[1/0.45]这个值。因此，这些公式现在也用于PAL、SECAM、576i和576p视频系统。

3.10.4 非CRT显示器

由于LCD、LCOS、DLP和plasma显示器并不是基于CRT的，因此它们具有不同的显示变换公式。为了简化与这些显示器的接口，它们的电子器件被设计为可以接收标准的gamma校准视频，然后对显示器面板的实际变换进行补偿。

3.11 恒定亮度问题

由于gamma和矩阵操作的错误顺序，U和V（或Cb和Cr）信号也对亮度信号（Y）有贡献。U和V的幅度不正确会导致接收到的亮度发生错误。这可能是由于带宽限制了U和V，或者U和V增益的非标称设置（颜色饱和）所致。

对于低的颜色频率，这没有问题。但对于高颜色频率，U和V将消失，这导致R′、G′和B′退

化为Y（只等于Y）。

参考文献

1. Benson. K. Blair, *Television Engineering Handbook*. McGraw-Hill, Inc., 1986.
2. Devereux, V. G., 1987, *Limiting of YUV digital video signals*, BBC Research Department Report BBC RD 1987 22.
3. EIA Standard EIA-189-A, July 1976, *Encoded Color Bar Signal*.
4. Faroudja, Yves Charles, *NTSC and Beyond. IEEE Transactions on Consumer Electronics*, Vol. 34, NO.1, February 1988.
5. IEC 61966-2-1, 1999, *Colour Management-Default RGB Colour Space-sRGB*.
6. IEC 61966-2-2, 2003, *Colour Management-Extended RGB Colour Space-scRGB*.
7. IEC 61966-2-4, 2006, *Colour Management-Gamut YCC Colour Space for Video Applications-xv YCC*.
8. ITU-R BT.470-6, 1998, *Conventional Television Systems*.
9. ITU-R BT.601-5, 1995, *Studio Encoding Parameters of Digital Television for Standard 4 : 3 and Widescreen 16 : 9 Aspect Radios*.
10. ITU-R BT.709-6, 2002, *Parameter Values for the HDTV Standards for Production and International Programme Exchange*.
11. Photo CD Information Bulletin, *Fully Utilizing Photo CD Images-PhotoYCC Color Encoding and Compression Schemes*, May 1994, Eastman Kodak Company.



第4章 视频信号概述

视频信号具有很多不同的选项，扫描线的数目、隔行还是逐行、模拟还是数字等。本章概要介绍了常用视频信号的格式及其时序。

4.1 数字分量视频背景

在数字分量视频（digital component video）中，视频信号是以数字形式（YCbCr或R′G′B′）表示的，仅当用于广播或录制时，才将其编码为复合NTSC、PAL或SECAM信号。

由于在576i PAL和SECAM制式之间交换视频资料遇到了困难，欧洲广播联盟（European Broadcasting Union，EBU）对数字分量视频产生了兴趣。因为，这种格式（数字分量视频）为解决这个问题带来了希望。那就是，不管信号源是来自使用PAL还是SECAM的国家，数字视频信号都可以是一样的，这就使得在后续操作中可以将其编码为合适的复合形式用于广播。运动图像与电视工程师协会（SMPTE）协商后，决定开发一种方法来支持国际节目交换（exchange），并将480i系统包括其中。

为了确定各种信号处理方法的质量和适用性，研究人员进行了一系列的实验演示。通过这些实验研究，确定了数字分量的编码、滤波和时序的主要参数，并将其编入了ITU-R BT.601标准中，因而BT.601成为了其他数字分量视频标准的起点。

4.1.1 编码范围

编码范围的选择需要综合考虑正常范围之外，信号所需容量和最小化量化失真（distortion）这两个要求。尽管视频信号中黑色的量化等级可以很好地进行合理定义，但白色的量化等级却容易受到由视频信号和设备限差（tolerance）引起的变化的影响。由滤波引起的噪声、增益变化和瞬变现象都可能导致超出正常范围的信号电平。

每个采样点需要8或10位用于每个YCbCr或R′G′B′分量。尽管8位编码会引入一些量化失真，但初步感觉是，大多数视频源包含足以淹没大多数量化失真的噪声。但是，如果视频源实际上是没有噪声的，那么量化失真就值得注意，因为在亮度渐进变化的区域会出现轮廓。另外，当在演播编辑环境中不同设备之间发送数据时，至少还需要额外的2个数据位用于表示YCbCr或R′G′B′颜色的小数部分，以减小舍入误差的影响。正是由于这些原因，大多数专业视频设备都使用10位表示YCbCr或R′G′B′颜色数据，准许2位YCbCr或R′G′B′颜色数据的小数部分得以保留。

起初的建议是让所有三个YCbCr颜色分量具有相等的编码范围。然而这些后来发生了变化，这种变化是为了使Y在白色处具有更大的空间用于表示量化等级，因为量化等级对白色的限制比黑色更容易被觉察。因此，标称的8位Y量化等级的表示范围为16~235，而标称的8位CbCr量化等级的表示范围为16~240（128对应于无色）。偶尔漂移到了其他量化等级中也是准许的，但绝不能在0或255处。

对于8位的系统，值0x00和0xFF被保留用于存放时序信息。而对于10位的系统，值0x000~0x003和0x3FC~0x3FF被保留用于保存时序信息，以维持与8位系统的兼容。

第3章讨论过用于生成75%彩条的YCbCr和R′G′B′颜色量化等级。数字R′G′B′信号定义为与Y

具有相同的标称量化等级，以提供更多的处理余地和简化R′G′B′和YCbCr之间的数字矩阵转换。

4.1.2 SDTV采样率的选择

对模拟R′G′B′或YUV视频信号实施行锁定（line-locked）采样。这种方法产生了一个静态的正交采样网格。在这个网格中，当前扫描线上的采样点落在前面扫描线和前面场的采样点的正下方，如图3-2至图3-11所示。

另一个更重要的特征是，采样在相位上是锁定的，这样采样点就位于模拟水平同步（0x0）信号下降沿50%幅度位置。这保证了不同的信号源生成的采样点在图像中处于相同的标称位置。使得不同的信号源都具有这一特性，可以简化从一个标准到另一个标准的转换。

对于480i和576i视频系统，起初考虑了几个Y采样频率，包括4倍F_{SC}采样率。然而，4倍F_{SC}采样率并不支持国际节目交换的要求。因此，为了支持一个单一的公共采样率，起初考虑的那些采样频率都被舍弃了。由于寻找可能的最低采样率（但仍然支持高质量的视频）是选择采样频率的一个目标，因此很长时间内都将12MHz采样率作为首选，但是最后考虑到该采样频率过于接近奈奎斯特频率极限，这会使滤波要求更加复杂。在对12~14.3MHz之间的频率进行考察后，可以很明显地知道，以13.5MHz的采样率对Y进行采样在480i和56i系统之间具有某些通用性。Cb和Cr为颜色微分信号，并不要求与Y具有同样的带宽，因此很多采样点都是以Y采样率的一半进行采样的，即6.75MHz。

38

现在普遍使用的4：2：2颜色标准起初是用于NTSC和PAL视频的，这意味着Y、U和V分别以颜色副载波频率的4倍、2倍和2倍频率进行采样。后来4：2：2表示被用于BT.601数字分量视频，这意味着Y、Cb和Cr的采样频率分为3.375MHz的4倍、2倍和2倍。现在4：2：2表示通常意味着Cb和Cr的采样率是Y的一半，而不管实际采用的采样率到底为多少。

使用13.5MHz进行采样，每行扫描线包含858个采样点（480i系统）或864个采样点（576i系统），扫描线由一个后面紧跟一个有效线周期（active line period）的数字消隐期（blanking interval）组成。在有效线周期期间，不管是480i还是576i系统，它们都使用720个采样点。每个有效线周期都具有相同的采样点数目，这简化了多标准设备和标准变换的设计。Cb和Cr采用6.75MHz的采样率（4：2：2采样），每个有效线周期包含360个Cr采样点和360个Cb采样点。

对于模拟系统，在进行重复处理时就出现了问题，会导致消隐期扩展、消隐边缘软化。在有效线周期内使用720个采样点，可以同时包容480i和576i制式的模拟消隐误差容限。因此，重复处理可以在不影响数字消隐期的情况下完成。用于定义模拟图像宽度的消隐仅需要进行一次即可，在显示器中或向上转换为模拟视频时是很合适的。

起初，BT.601仅支持具有4：3宽高比的480i和576i系统（有效分辨率为720×480i和720×576i）。对16：9宽高比的支持是通过使用18MHz采样率来实现的（有效分辨率为960×480i和960×576i）。

4.1.3 EDTV采样率的选择

ITU BT.1358定义了逐行SDTV视频信号，也就是我们所知的480p或576p，或增强高清晰度电视（EDTV）。为了保持与BT.601使用同样的静态正交采样网格，ITU BT.1358使用的采样率加倍了，变为27MHz（4：3的宽高比）或36MHz（16：9的宽高比）。

4.1.4 HDTV采样率的选择

ITU BT.709分别定义了720p、1080i和1080p视频信号。对于HDTV，使用了不同的技术，每线的有效采样点数量和每帧的有效线数目为恒定的，不考虑采用什么样的帧率。因此，为了保持静

态的正交采样网格，每种帧率都使用不同的采样时钟速率。

4.2 480i和480p系统

4.2.1 隔行模拟复合视频

（M）NTSC和（M）PAL为模拟复合视频信号，它们在一个单一信号中包含着所有的颜色和时序信息。这些模拟接口使用每帧525线的分辨率，将在第8章对此进行详细讨论。

4.2.2 隔行模拟分量视频

模拟分量信号由三种信号组成，这些信号可以是模拟R'G'B'或YPbPr。至于480i［具有典型的每帧480有效扫描线并且是隔行（interlaced）扫描的］，帧率通常是29.97Hz（30/1.001），这是为了与（M）NTSC的时序兼容。模拟接口使用每帧525线，在23~262和286~525线上显示有效视频，如图4-1所示。

39

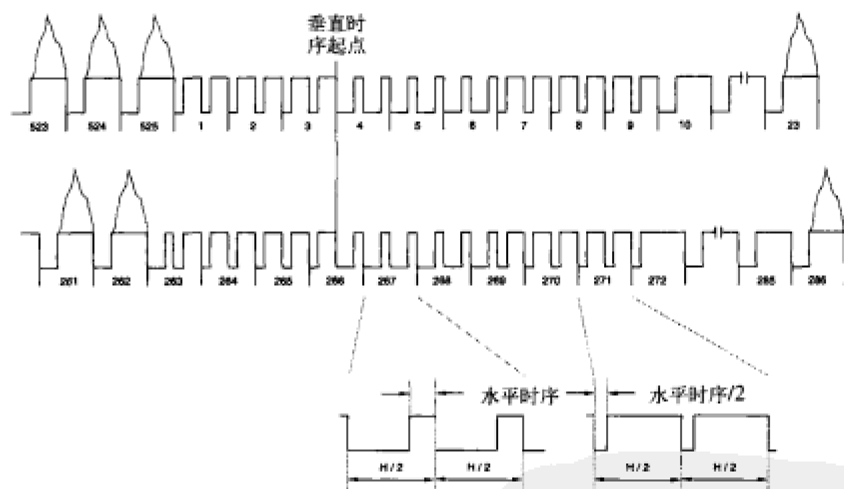


图4-1 480i垂直间隔时序

对于29.97Hz的帧率，每行扫描线时间（H）约为63.556 μs。详细的水平时序取决于所用的特定视频接口，将在第5章对此进行讨论。

4.2.3 逐行模拟分量视频

模拟分量信号由三个信号组成，这些信号为模拟R'G'B'或YPbPr信号。至于480p［具有典型的每帧480有效扫描线并且是逐行（progressive）扫描的］，帧率通常是59.94Hz（60/1.001），这是为了更易于与（M）NTSC的时序兼容。模拟接口使用每帧525线，在45~524线上显示有效视频，如图4-2所示。



图4-2 480p垂直间隔时序

对于59.94Hz的帧率，每行扫描线时间（H）大约为31.776 μs。详细的水平时序取决于所使用的特定视频接口，将在第5章对此进行讨论。

4.2.4　隔行数字分量视频

BT.601和SMPTE 267M规定了480i数字R'G'B'或YCbCr视频信号的表示方式。BT.601和SMPTE 267M定义的有效分辨率，其1×Y和R'G'B'采样率（F_S），以及帧率分别为：

960×480i　18.0MHz　29.97Hz
720×480i　13.5MHz　29.97Hz

对于其他通用有效分辨率，其1倍采样率（F_S）和帧率分别为：

864×480i　16.38MHz　29.97Hz
704×480i　13.50MHz　29.97Hz
640×480i　12.27MHz　29.97Hz
544×480i　10.12MHz　29.97Hz
528×480i　9.900MHz　29.97Hz
480×480i　9.000MHz　29.97Hz
352×480i　6.750MHz　29.97Hz

40
↕
41

864×480i为16：9方形像素格式，而640×480i为4：3的方形像素格式。尽管理想的16：9分辨率为854×480i，但864×480i可以支持MPEG 16×16块结构。704×480i格式是通过720×480i格式来完成的，对每个有效扫描线的前8个和后8个采样点进行了消隐处理。模拟和数字信号之间的关系见图4-3~图4-7的例子。

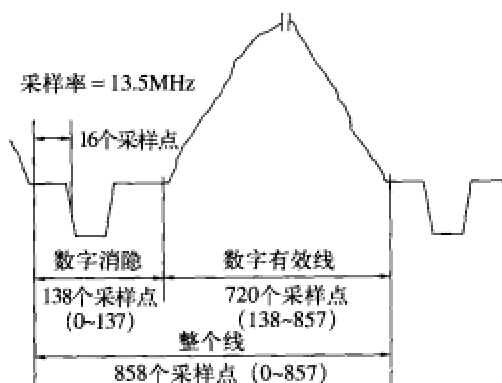


图4-3　480i模－数关系（4：3宽高比、29.97Hz的帧率、13.5MHz的采样时钟）。BT.601规定前肩为16个采样点；CEA-861D（DVI和HDMI时序）规定前肩为19个采样点

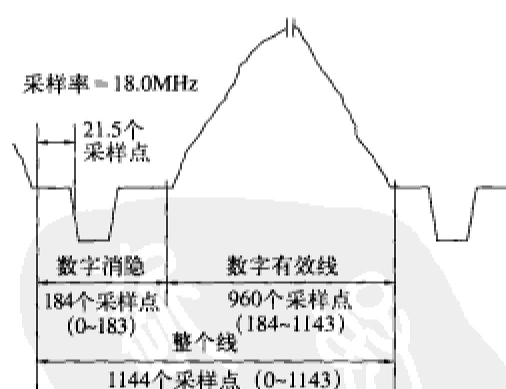


图4-4　480i模－数关系
（16：9宽高比、29.97Hz的帧率、18MHz的采样时钟）

H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-8。其中的H、V和F时序与视频压缩标准兼容，而与第6章要讨论的BT.656并不兼容。

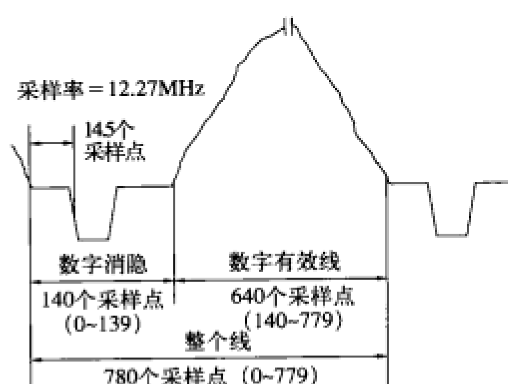


图4-5 480i模－数关系

（4：3宽高比、29.97Hz的帧率、12.27MHz的采样时钟）

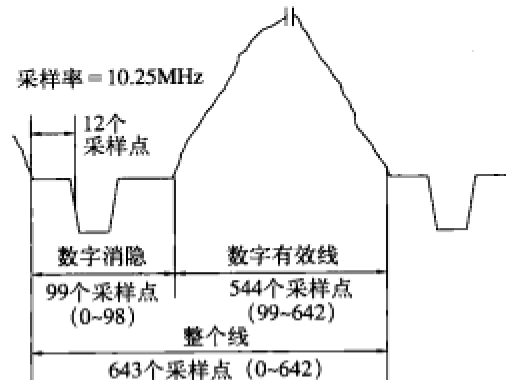


图4-6 480i模－数关系

（4：3宽高比、29.97Hz的帧率、10.125MHz的采样时钟）

42

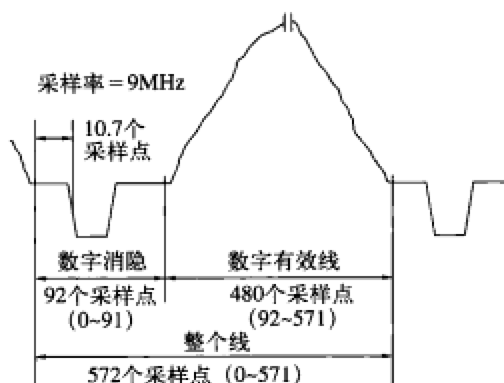
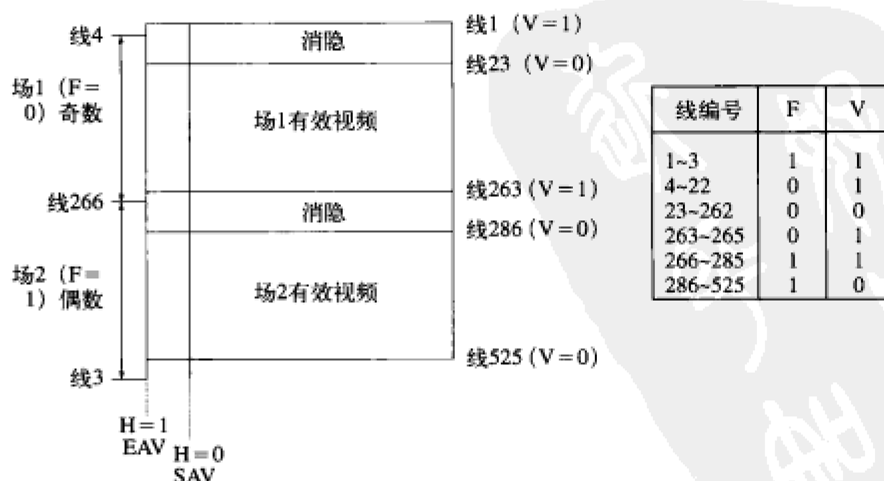


图4-7 480i模－数关系

（4：3宽高比、29.97Hz的帧率、9MHz的采样时钟）

43



线编号	F	V
1~3	1	1
4~22	0	1
23~262	0	0
263~265	0	1
266~285	1	1
286~525	1	0

图4-8 480i数字垂直时序（480有效线）。在数字线的开始处，F和V按照EAV顺序改变状态。注意，数字线数目在水平时序开始之前改变状态，如图4-3~图4-7所示。这些有效线被SMPTE RP-202、ATSC A/54a和ARIB STD-B32标准所采用。CEA-861D（DVI和HDMI时序）规定有效视频为线22~261和285~524。IEC61834-2、ITU-R BT.1618和SMPTE 314M（DV格式）规定有效视频线为23~262和285~524。ITU-R BT.656规定有效视频为线20~263和283~525，因此每帧总的有效线为487

44

4.2.5 逐行数字分量视频

BT.1358和SMPTE 293M规定了480p数字R′G′B′或YCbCr视频信号的表示方式。BT.1358和SMPTE 293M定义的有效分辨率，其1倍采样率（F_S）和帧率为：

960×480p 36.0MHz 59.94Hz

720×480p 27.0MHz 59.94Hz

对于其他通用有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

864×480p 32.75MHz 59.94Hz

704×480p 27.00MHz 59.94Hz

640×480p 24.54MHz 59.94Hz

544×480p 20.25MHz 59.94Hz

528×480p 19.80MHz 59.94Hz

480×480p 18.00MHz 59.94Hz

352×480p 13.50MHz 59.94Hz

864×480p为16：9方形像素格式，而640×480p为4：3的方形像素格式。尽管理想的16：9分辨率为854×480p，但864×480p可以支持MPEG 16×16块结构。704×480p格式是通过720×480p格式来完成的，对每个有效扫描线的前8个和后8个采样点进行了消隐处理。模拟和数字信号之间的关系的例子见图4-9~图4-12。

45

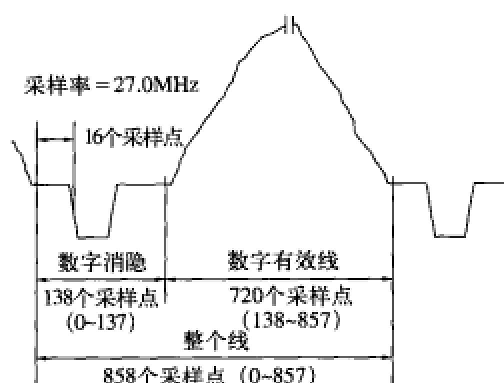


图4-9 480p模－数关系

（4：3宽高比、59.94Hz的帧率、27MHz的采样时钟）

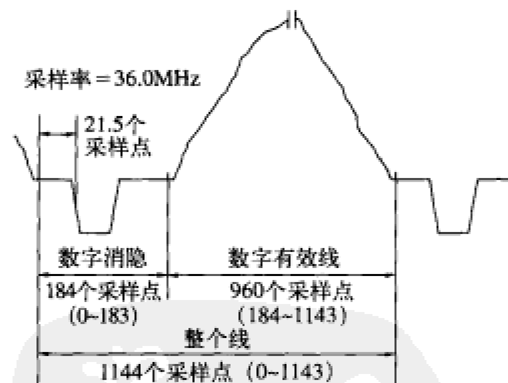


图4-10 480p模－数关系

（16：9宽高比、59.94Hz的帧率、36MHz的采样时钟）

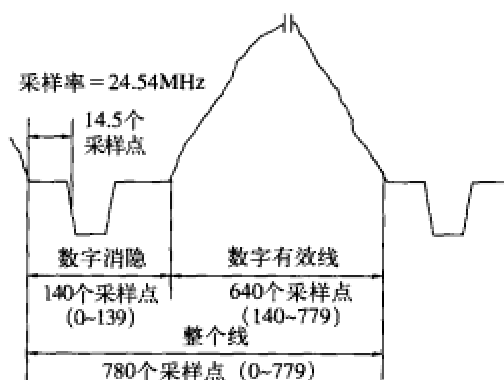


图4-11 480p模－数关系

（4：3宽高比、59.94Hz的帧率、24.54MHz的采样时钟）

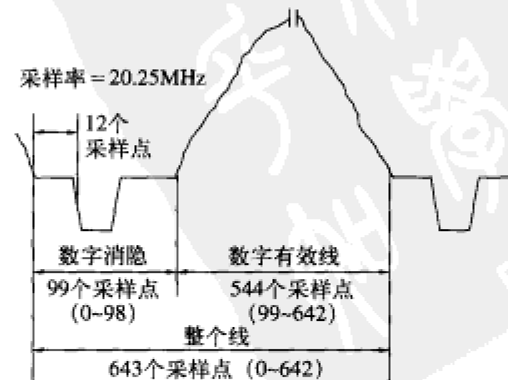
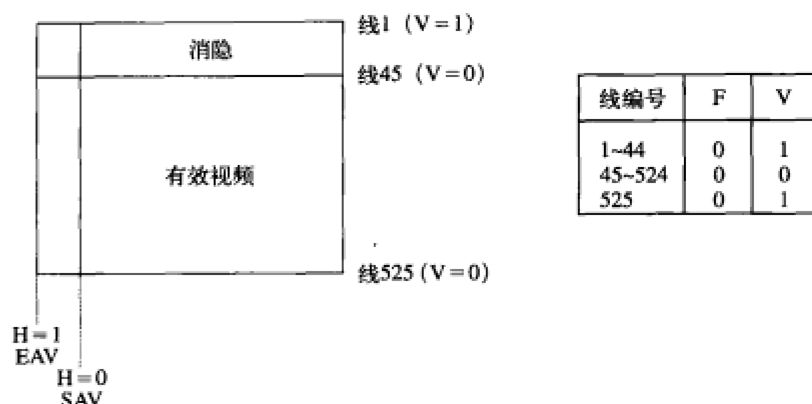


图4-12 480p模－数关系

（4：3宽高比、59.94Hz的帧率、20.25MHz的采样时钟）

H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-13。其中的H、V和F时序与视频压缩标准兼容，而与第6章要讨论的BT.656并不兼容。



线编号	F	V
1~44	0	1
45~524	0	0
525	0	1

图4-13 480p数字垂直时序（480有效线）。V在数字线的开始处按照EAV顺序改变状态。注意，数字线编号在水平时序开始之前改变状态，如图4-9~图4-12所示。这些有效线被SMPTE RP-202、ATSC A/54和ARIB STD-B32标准所采用。CEA-861（DVI和HDMI时序）规定有效视频为线43~522

46~47

4.2.6 SIF和QSIF

SIF定义的有效分辨率为352×240p。方形像素SIF定义的有效分辨率为320×240p。

QSIF定义的有效分辨率为176×120p。方形像素QSIF定义的效分辨率为160×120p。

4.3 576i和576p系统

4.3.1 隔行模拟复合视频

（B、D、G、H、I、N、N_C）PAL为模拟复合视频信号，它们在一个单一信号中包含了所有的颜色和时序信息。这些模拟接口使用每帧625线的分辨率，将在第8章对此进行详细讨论。

4.3.2 隔行模拟分量视频

模拟分量信号由三种信号组成，这些信号可以是模拟R'G'B'或YPbPr。至于576i［具有典型的每帧576有效扫描线并且是隔行（interlaced）扫描的］，帧率通常是25Hz，这是为了与PAL的时序兼容。模拟接口使用每帧625线，在23~310和336~623线上显示有效视频，如图4-14所示。

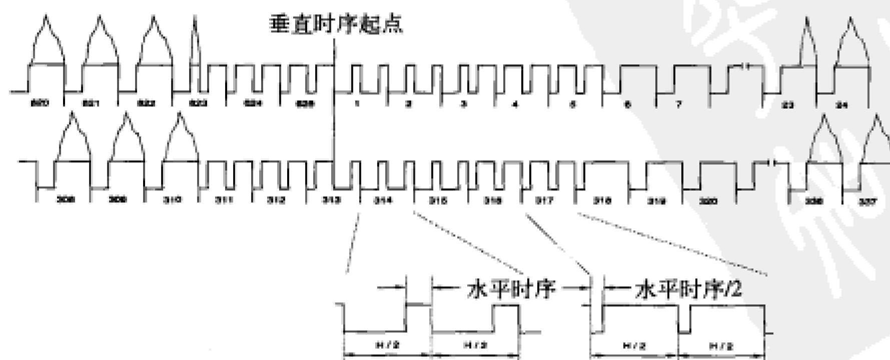


图4-14 576i垂直间隔时序

对于25Hz的帧率，每个扫描线时间（H）约为64 μs。详细的水平时序取决于所用的特定视频接口，将在第5章对此进行讨论。

4.3.3 逐行模拟分量视频

模拟分量信号由三个信号组成，这些信号为模拟R′G′B′或YPbPr信号。至于576p［具有典型的每帧576有效扫描线，并且是逐行扫描的］，帧率通常是50Hz，这是为了更易于与PAL时序兼容。模拟接口使用每帧625线，在45~620线上显示有效视频，如图4-15所示。

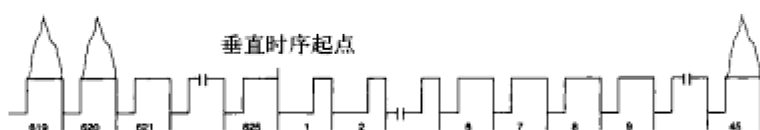


图4-15 576p垂直间隔时序

对于50Hz的帧率，每个扫描线时间（H）大约为32微秒。详细的水平时序依赖于所使用的特定视频接口，将在第5章对此进行讨论。

4.3.4 隔行数字分量视频

BT.601规定了576i数字R′G′B′或YCbCr视频信号的表示方式。BT.601和SMPTE 267M定义的有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

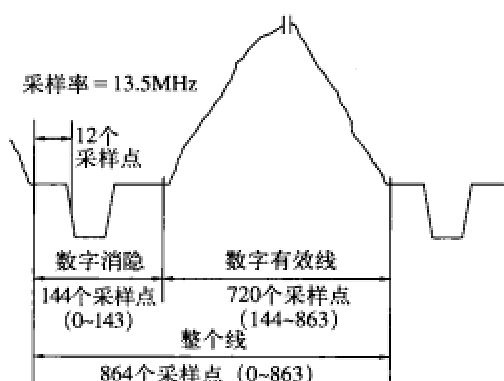
960×576i 18.0MHz 25Hz
720×576i 13.5MHz 25Hz

对于其他通用有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

1024×576i 19.67MHz 25Hz
768×576i 14.75MHz 25Hz
704×576i 13.50MHz 25Hz
544×576i 10.12MHz 25Hz
480×576i 9.00MHz 25Hz

48~49

1024×576i为16：9的方形像素格式，而768×576i为4：3的方形像素格式。704×576p格式是通过720×576p格式来完成的，对每个有效扫描线的前8个和后8个采样点进行了消隐处理。模拟和数字信号之间的关系见图4-16~图4-19的例子。



50

图4-16 576i模－数关系
（4：3宽高比、25Hz的帧率、13.5MHz的采样时钟）

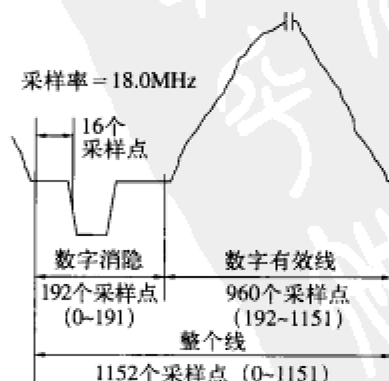


图4-17 576i模－数关系
（16：9宽高比、25Hz的帧率、18MHz的采样时钟）

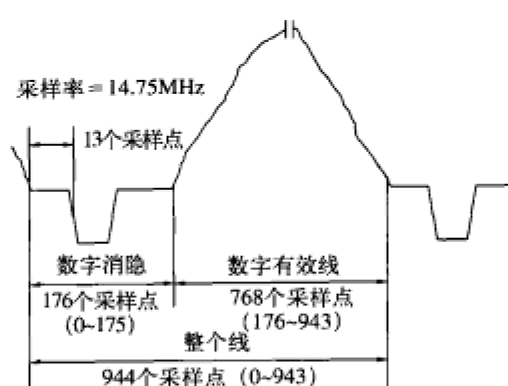


图4-18 576i模－数关系
（4：3宽高比、25Hz的帧率、14.75MHz的采样时钟）

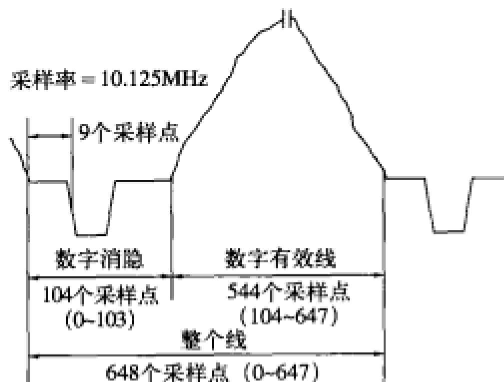


图4-19 576i模－数关系
（4：3宽高比、25Hz的帧率、10.125MHz的采样时钟）

51

H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-20。其中的H、V和F时序与视频压缩标准兼容，而与第6章要讨论的BT.656并不兼容。



线编号	F	V
1~22	0	1
23~310	0	0
311~312	0	1
313~335	1	1
336~623	1	0
624~625	1	1

图4-20 576p数字垂直时序（576行有效线）。在数字线的开始处，V按照EAV顺序改变状态。注意，数字线数目在水平时序开始之前改变状态，如图4-16~图4-19所示。IEC 61834-2、ITU-R BT.1618和SMPTE 314M（DV格式）规定有效视频为线23~310和335~622

52

4.3.5 逐行数字分量视频

BT.1358规定了576p数字R′G′B′或YCbCr视频信号的表示方式。BT.1358定义的有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

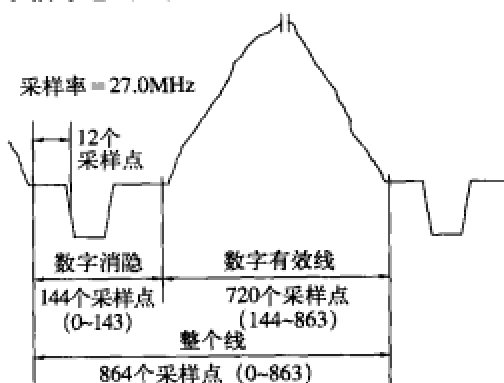
960×576p　36.0MHz　50Hz
720×576p　27.0MHz　50Hz

对于其他通用有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

1024×576p　39.33MHz　50Hz
768×576p　29.5MHz　50Hz
704×576p　27.00MHz　50Hz
544×576p　20.25MHz　50Hz

480×576p 18.00MHz 50Hz

1024×576p为16：9方形像素格式，而768×576p为4：3的方形像素格式。704×576p格式是通过720×576p格式来完成的，对每个有效扫描线的前8个和后8个采样点进行了消隐处理。模拟和数字信号之间的关系的例子见图4-21~图4-24。



53

图4-21 576p模－数关系

（4：3宽高比、50Hz的帧率、27MHz的采样时钟）

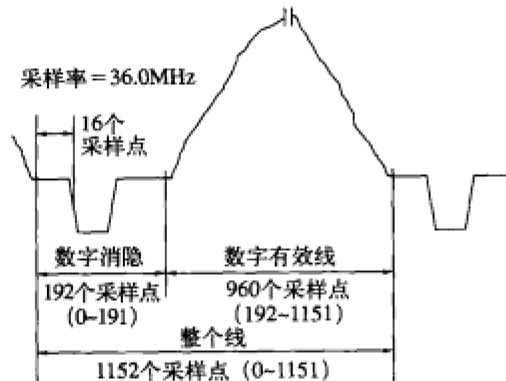
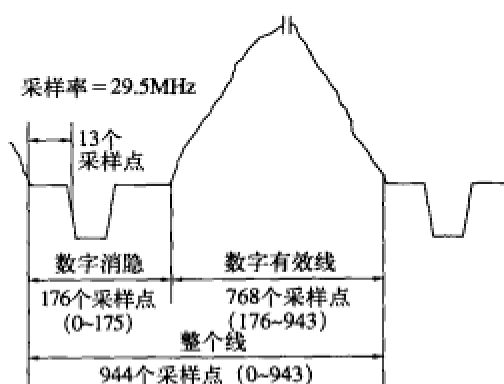


图4-22 576p模－数关系

（16：9宽高比、50Hz的帧率、36MHz的采样时钟）



54

图4-23 576p模－数关系

（4：3宽高比、50Hz的帧率、29.5MHz的采样时钟）

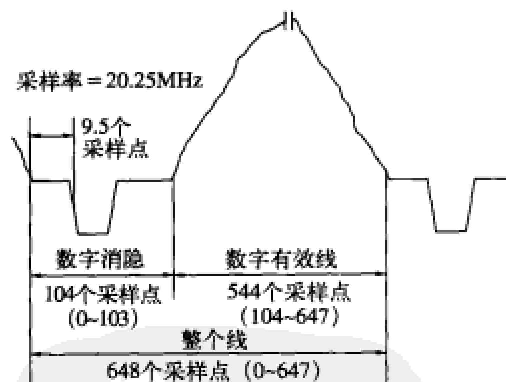
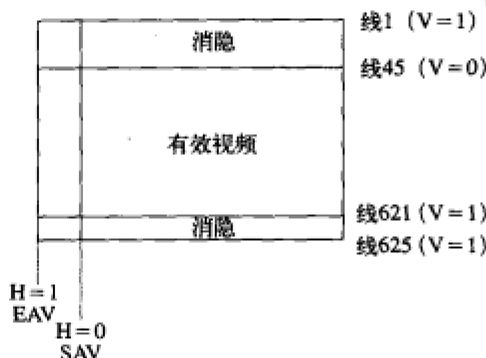


图4-24 576p模－数关系

（4：3宽高比、50Hz的帧率、20.25MHz的采样时钟）

H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-25。其中的H、V和F时序与视频压缩标准兼容，而与第6章要讨论的BT.656并不兼容。



线编号	F	V
1~44	0	1
45~620	0	0
621~625	0	1

55

图4-25 576p数字垂直时序（576条有效线）。在数字线的开始处，V按照EAV顺序改变状态。注意，数字线编号在水平时序开始之前改变状态，如图4-21~图4-24所示

4.4　720p系统

4.4.1　逐行模拟分量视频

模拟分量信号由三个信号组成，这些信号为模拟R'G'B'或YPbPr信号。至于720p［具有典型的每帧720有效扫描线并且是逐行（progressive）扫描的］，帧率通常是59.94Hz（60/1.001），这是为了简化（M）NTSC视频的生成。模拟接口每帧使用750线，在26~745线上显示有效视频，如图4-26所示。



图4-26　720p垂直间隔时序

对于59.94Hz的帧率，每个扫描线时间（H）大约为22.24μs。详细的水平时序取决于所使用的特定视频接口，将在第5章对此进行讨论。

4.4.2　逐行数字分量视频

SMPTE 296M规定了720p数字R'G'B'或YCbCr视频信号的表示方式。SMPTE 296M定义的有效分辨率，其1×Y和R'G'B'采样率（F_S）以及帧率分别为：

1280×720p	74.176MHz	23.976Hz
1280×720p	74.250MHz	24.000Hz
1280×720p	74.250MHz	25.000Hz
1280×720p	74.176MHz	29.970Hz
1280×720p	74.250MHz	30.000Hz
1280×720p	74.250MHz	50.000Hz
1280×720p	74.176MHz	59.940Hz
1280×720p	74.250MHz	60.000Hz

注意，1280×720p使用的是方形像素和16：9的宽高比。模拟和数字信号之间的关系的例子见图4-27、图4-28和表4-1。H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-29。 56

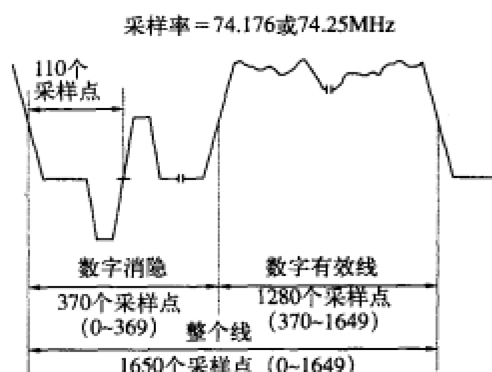


图4-27　720p模－数关系（宽高比为16：9、帧率为59.94 Hz、采样时钟为74.176MHz，以及帧率为60Hz、采样时钟为74.25MHz）

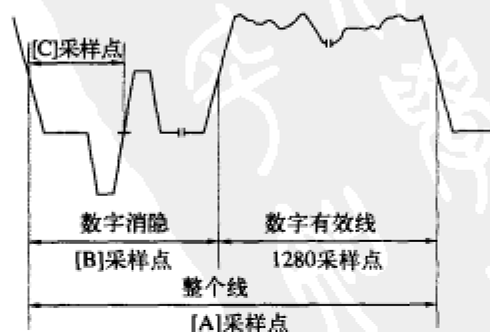


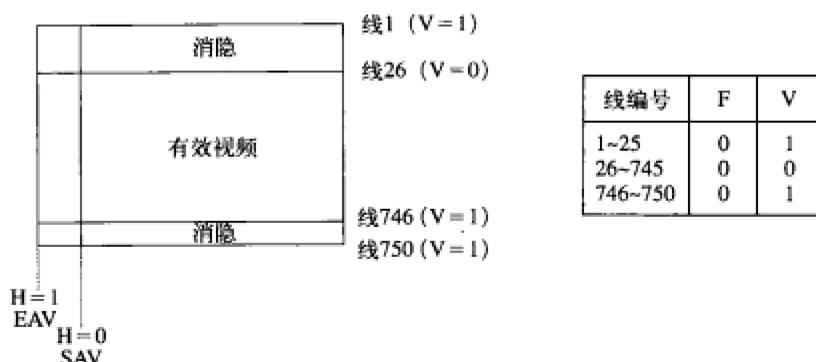
图4-28　一般720p模－数关系

57

表4-1 图4-28中各种720p模－数参数

有效水平采样点	帧率（Hz）	1×Y采样率（MHz）	总的水平采样点（A）	水平消隐采样点（B）	C采样点
1280	24/1.001	74.25/1.001	4125	2845	2585
	24	74.25	4125	2845	2585
	25[1]	48	1536	256	21
	25[1]	49.5	1584	304	25
	25	74.25	3960	2680	2420
	30/1.001	74.25/1.001	3300	2020	1760
	30	74.25	3300	2020	1760
	50	74.25	1980	700	440
	60/1.001	74.25/1.001	1650	370	110
	60	74.25	1650	370	110

1. 基于31.250kHz的线频率对于基于CRT、50Hz的HDTV是有益的。时序脉冲为－300mV的双极信号，而不是±300mV的三极信号。720p的视频内容在垂直方向上缩放为1152i有效扫描线，总的扫描线为1250i，而不是750p。



线编号	F	V
1~25	0	1
26~745	0	0
746~750	0	1

图4-29 720p数字垂直时序（720有效线）。在数字线的开始处，V按照EAV顺序改变状态。注意，数字线在水平时序之前改变状态，如图4-27和图4-28所示

58

4.5 1080i和1080p系统

4.5.1 隔行模拟分量视频

模拟分量信号由三个信号组成，这些信号可以是模拟R′G′B′或YPbPr。至于1080i［具有典型的每帧1080有效扫描线并且是隔行（interlaced）扫描的］，帧率通常是25Hz或29.97Hz（30/1.001），这简化了（B、D、G、H、I）PAL或（M）NTSC视频的生成。模拟接口使用每帧1125线，在21~560和584~1123线上显示有效视频，如图4-30所示。

MPEG-2和MPEG-4系统使用1088线，而不是使用1080，这是为了每帧具有32的整倍数扫描线。在这种情况下，每个场在有效视频之后使用了额外的4线。

对于25Hz的帧率，每行扫描线的时间（H）约为35.56 μs。对于29.97Hz的帧率，每行扫描线的时间约为29.66 μs。详细的水平时序依赖于所用的特定视频接口，将在第5章对此进行讨论。

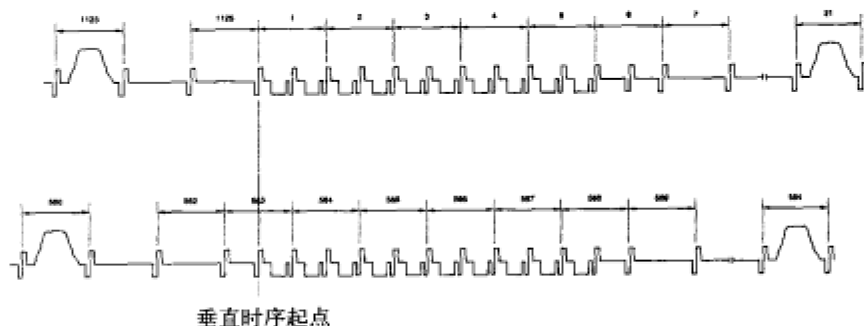


图4-30 1080i垂直间隔时序

4.5.2 1152i格式

1152i（总数为2350）有效线格式并不是广播发射格式。然而，它却被用于从HD机顶盒和DVD播放机直到50Hz基于CRT的HDTV之间的各种模拟互连标准，使得50Hz的HDTV的水平频率可以使用固定的31.25kHz，从而降低了费用。其他的HDTV显示器技术，如DLP、LCD和等离子显示器，能够处理720p50（带VBI的750p50）和1080i25（带VBI的1125i25）模拟信号的本地时序。

机顶盒或DVD播放器将720p50和1080i25的视频内容转换为1152i25格式。将1280×720p50的视频内容缩放为1280×1152i25格式；将1920×1080i25以1920×1152i25的宽银幕模式提供给外界。为了校准1080i25的几何参数（这可以通过垂直时序信号对其进行识别），HDTV将具有一个标称垂直缩放模式。

4.5.3 逐行模拟分量视频

模拟分量信号由三个信号组成，这些信号为模拟R′G′B′或YPbPr信号。至于1080p［具有典型的每帧1080行有效扫描线并且是逐行（progressive）扫描的］，帧率通常是50或59.94Hz（60/1.001），这是为了简化（B、D、G、H、I）PAL或（M）NTSC视频信号的生成。模拟接口使用每帧1125线，在42~1121线上显示有效视频，如图4-31所示。



图4-31 1080p垂直间隔时序

为了使每帧视频的扫描线数是16的整倍数，MPEG-2和MPEG-4系统使用1088线，而不是1080线。为此，在每场视频数据的有效视频之后需要增加额外的8行扫描线。

对于50Hz的帧率，每行扫描线时间大约为17.78 μs。对于59.94Hz的帧率，每行扫描线时间大约为14.83 μs。详细的水平时序取决于所使用的特定视频接口，将在第5章对此进行讨论。 59

4.5.4 隔行数字分量视频

ITU-R BT.709和SMPTE 274M规定了1080i数字R′G′B′或YCbCr视频信号的数字分量格式。BT.709和SMPTE 267M定义的有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

1920×1080i 74.250MHz 25.00Hz

1920×1080i 74.176MHz 29.97Hz

1920 × 1080i　74.250MHz　30.00Hz

注意，ITU-R BT.709和SMPTE 274M使用的是方形像素和16：9的宽高比。对于其他通用有效分辨率，其1 × Y和R′G′B′采样率（F_S）及帧率分别为：

1280 × 1080i　49.500MHz　25.00Hz

1280 × 1080i　49.451MHz　29.97Hz

1280 × 1080i　49.500MHz　30.00Hz

1440 × 1080i　55.688MHz　25.00Hz

1440 × 1080i　55.632MHz　29.97Hz

1440 × 1080i　55.688MHz　30.00Hz

60 ↓ 62

模拟和数字信号之间的关系的例子见图4-32、图4-33和表4-2。H（水平消隐）和V（垂直消隐）信号的定义见图4-34。

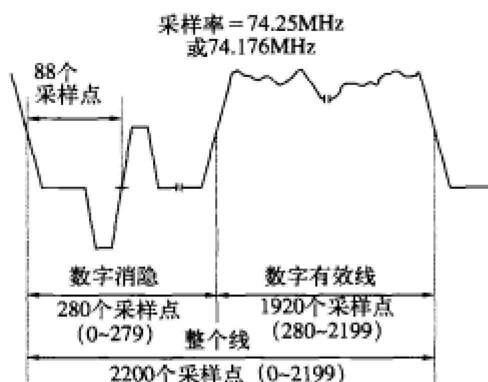


图4-32　1080i模－数关系（16：9宽高比、29.97Hz的帧率、74.176MHz的采样时钟，以及30Hz的帧率、74.25MHz的采样时钟）

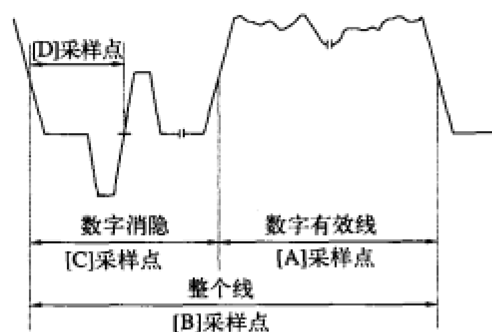
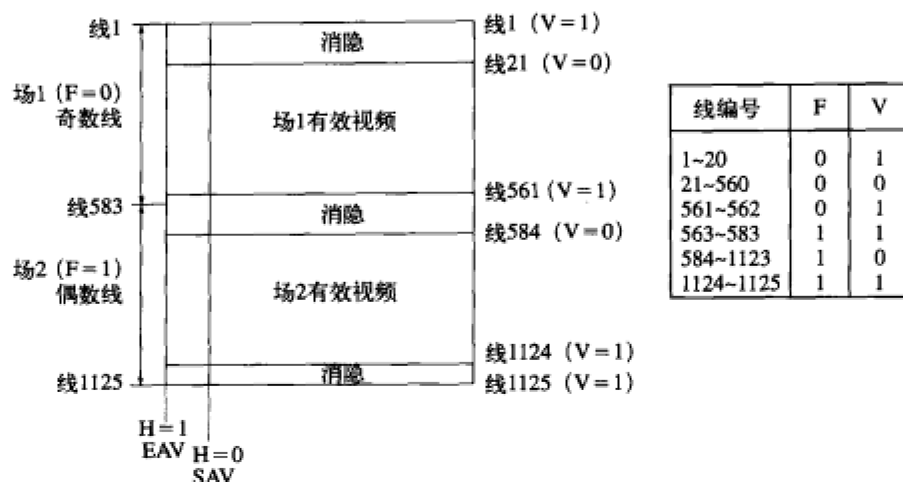


图4-33　一般的1080i模－数关系

表4-2　图4-33中的各种1080i模－数参数

有效水平采样点（A）	帧率（Hz）	1 × Y采样率（MHz）	总的水平采样点（B）	水平消隐采样点（C）	D采样点
1920	25[1]	72	2304	384	32
	25[1]	74.25	2376	456	38
	25	74.25	2640	720	528
	30/1.001	74.25/1.001	2200	280	88
	30	74.25	2200	280	88
1440	25[1]	54	1728	288	24
	25	55.6875	1980	540	396
	30/1.001	55.6875/1.001	1650	210	66
	30	55.6875	1650	210	66
1280	25[1]	48	1536	256	21
	25	49.5	1760	480	352
	30/1.001	49.5/1.001	1466.7	186.7	58.7
	30	49.5	1466.7	186.7	58.7

1. 基于31.250kHz的线频率对于基于CRT、50Hz的HDTV是有用的。时序脉冲为－300mV的双极信号，而不是±300mV的三极信号。1080i的视频内容在垂直方向上缩放为1152i有效扫描线，总的扫描线为1250i，而不是1125i。



线编号	F	V
1~20	0	1
21~560	0	0
561~562	0	1
563~583	1	1
584~1123	1	0
1124~1125	1	1

图4-34 1080i数字垂直时序（1080有效线）。在数字线的开始处，V按照EAV顺序改变状态。注意，数字线在水平时序之前改变状态，如图4-32和图4-33所示

63

4.5.5 逐行数字分量视频

ITU-R BT.709和SMPTE 274M规定了1080p数字R′G′B′或YCbCr的数字分量格式。BT.709和SMPTE 267M定义的有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

1920×1080p 74.176MHz 23.976Hz
1920×1080p 74.250MHz 24.000Hz
1920×1080p 74.250MHz 25.000Hz
1920×1080p 74.176MHz 29.970Hz
1920×1080p 74.250MHz 30.000Hz
1920×1080p 148.50MHz 50.000Hz
1920×1080p 148.35MHz 59.940Hz
1920×1080p 148.50MHz 60.000Hz

注意，BT.709和SMPTE 267M使用方形像素和16：9宽高比。对于其他通用有效分辨率，其1×Y和R′G′B′采样率（F_S）以及帧率分别为：

1280×1080p 49.451MHz 23.976Hz
1280×1080p 49.500MHz 24.000Hz
1280×1080p 49.500MHz 25.000Hz
1280×1080p 49.451MHz 29.970Hz
1280×1080p 49.500MHz 30.000Hz
1280×1080p 99.000MHz 50.000Hz
1280×1080p 98.901MHz 59.940Hz
1280×1080p 99.000MHz 60.000Hz
1440×1080p 55.632MHz 23.976Hz
1440×1080p 55.688MHz 24.000Hz
1440×1080p 55.688MHz 25.000Hz
1440×1080p 55.632MHz 29.970Hz
1440×1080p 55.688MHz 30.000Hz
1440×1080p 111.38MHz 50.000Hz

1440 × 1080p　111.26MHz　59.940Hz
1440 × 1080p　111.38MHz　60.000Hz

模拟和数字信号之间关系的例子见图4-35、图4-36以及表4-3。H（水平消隐）、V（垂直消隐）和F（场）信号的定义见图4-37。

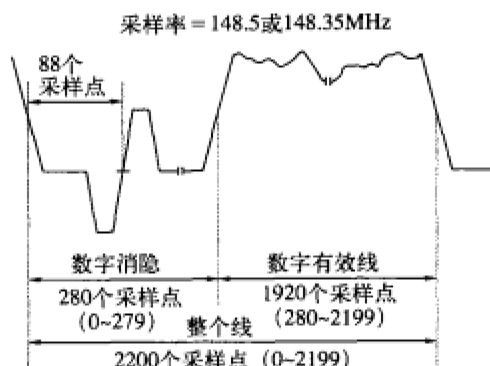


图4-35　1080p模－数关系（宽高比为16：9、帧率为59.94Hz、采样时钟为148.35MHz，以及帧率为60Hz、采样时钟为148.5MHz）

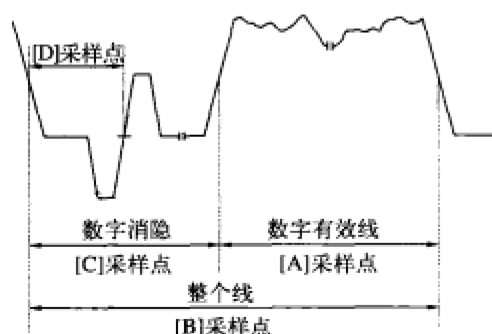
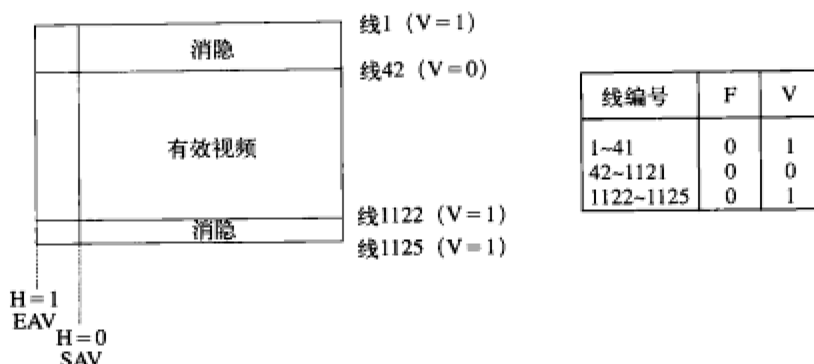


图4-36　一般的1080p模－数关系

表4-3　图4-36中的各种1080p模－数参数

有效水平采样点（A）	帧率（Hz）	1 × Y采样率（MHz）	总的水平采样点（B）	水平消隐采样点（C）	D采样点
1920	24/1.001	74.25/1.001	2750	830	638
	24	74.25	2750	830	638
	25	74.25	2640	720	528
	30/1.001	74.25/1.001	2200	280	88
	30	74.25	2200	280	88
	50	148.5	2640	720	528
	60/1.001	148.5/1.001	2200	280	88
	60	148.5	2200	280	88
1440	24/1.001	55.6875/1.001	2062.5	622.5	478.5
	24	55.6875	2062.5	622.5	478.5
	25	55.6875	1980	540	396
	30/1.001	55.6875/1.001	1650	210	66
	30	55.6875	1650	210	66
	50	111.375	1980	540	396
	60/1.001	111.375/1.001	1650	210	66
	60	111.375	1650	210	66
1280	24/1.001	49.5/1.001	1833.3	553.3	425.3
	24	49.5	1833.3	553.3	425.3
	25	49.5	1760	480	352
	30/1.001	49.5/1.001	1466.7	186.7	58.7
	30	49.5	1466.7	186.7	58.7
	50	99	1760	480	352
	60/1.001	99/1.001	1466.7	186.7	58.7
	60	99	1466.7	186.7	58.7



线编号	F	V
1~41	0	1
42~1121	0	0
1122~1125	0	1

图4-37　1080p数字垂直时序（1080有效线）。在数字线的开始处，V按照EAV顺序改变状态。注意，数字线在水平时序之前改变状态，如图4-35和图4-36所示

4.6　其他视频系统

有些消费类显示器已经采用其他的分辨率作为其固有分辨率，如基于LCD和等离子技术的显示器。常用的有效分辨率及其名称分别为：

640 × 400	VGA
640 × 480	VGA
854 × 480	WVGA
800 × 600	SVGA
1024 × 768	XGA
1280 × 768	WXGA
1366 × 768	WXGA
1024 × 1024	XGA
1280 × 1024	SXGA
1600 × 1024	WSXGA
1600 × 1200	UXGA
1920 × 1200	WUXGA

这些分辨率及其时序是由视频电子标准协会（Video Electronics Standards Association，VESA）为计算机显示器定义的。基于这些固有分辨率的显示器通常可以接收很多输入分辨率，它是通过将视频源进行缩放来实现与显示器分辨率的匹配的。

参考文献

1. CEA-861D, *A DTV Profile for Uncompressed High Speed Digital Interfaces*, July 2006.
2. EIA-770.1, *Analog 525-line Component Video Interface-Three Channels*, November 2001.
3. EIA-770.2, *Standard-Definition TV Analog Component Video Interface*, November 2001.
4. EIA-770.3, *High-Definition TV Analog Component Video Interface*, November 2001.
5. ITU-R BT.601-5, 1995, *Studio Encoding of Digital Television for Standard 4：3 and Widescreen 16：9 Aspect Radios.*
6. ITU-R BT-709-5, 2002, *Parameter Values for the HDTV Standards for Production and International Programme Exchange.*
7. ITU-R BT.1358, 1998, *Studio Parameters of 625 and 525 Line Progressive Scan Television Systems.*
8. SMPTE 267M-1995, *Television-Bit-Parallel Digital Interface—Component Video Signal 4：2：2 16 × 9 Aspect Radios.*
9. SMPTE 274M-2005, *Television-1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates.*
10. SMPTE 293M-2003, *Television-720 × 483 Active Line at 59.94Hz Progressive Scan Production—Digital Representation.*
11. SMPTE 296M-2001, *Television—1280 × 720 Progressive Image Sample Structure, Analog and Digital Representation and Analog Interface.*

第5章 模拟视频接口

多年以来，消费类市场上使用的主要视频信号一直是复合NTSC或者PAL视频（见图8-2和图8-13）。人们也试图进一步支持S-Video，但是直到最近，它还仅限于S-VHS VCR和高端电视机。

随着DVD播放机、数字机顶盒和DTV的出现，人们对在消费类市场提供高质量视频重新产生了兴趣。这些设备不仅支持很高质量的复合和S-Video信号，而且还准许选择使用模拟R′G′B′或YPbPr视频。

使用模拟R′G′B′或YPbPr视频去除了NTSC/PAL编码和解码过程中产生的人为噪声（artifact）。因此，图像更清晰、噪声更小。拥有更多的色度带宽，增加了水平清晰度。

5.1 S-Video接口

RCA梅花连接头（消费类市场）或BNC连接头（专业视频市场）对复合NTSC或PAL视频信号进行转换，其方法是将亮度（Y）和色度（C）视频信号加在一起。然后，为了显示图像，电视就需要将这些Y和C视频信号分离开。问题是，Y/C分离过程总是不完美的，这个问题将在第9章中讨论。

目前，许多视频分量支持4针的“S1”S-Video连接头，如图5-1所示（母连接头视图）。该连接头保持亮度（Y）和色度（C）视频信号分离，省去在TV上进行Y/C分离的过程。结果是图像更清晰，并具有更少的噪声。图9-2和图9-3表示的是Y信号，图9-10和图9-11表示的是C信号。

NTSC和PAL VBI（场消隐期）数据可以表示为480i或576i视频信号，将在第8章中对此进行讨论。

当提供的是宽屏（16：9）失真节目（水平压缩25%）时，“S2”版在C信号中增加了+5V DC偏移。16：9的TV检测DC偏移并在水平方向上对4：3图像进行扩展以填充屏幕，恢复节目的正确宽高比。当节目使用“Letterbox模式”后，“S3”版也支持+2.3V的偏移。

IEC 60993-5标准规定了S-Video的连接头和信号电平。

扩展S-Video接口

PC市场也使用扩展S-Video接口。这些接口有7针，如图5-1所示，并与4针接口兼容。

其他3针的使用是由制造商来定义的。它们可以用于支持I²C接口（SDA双向数据针和SCL时钟针）、+12V电源、复合NTSC/PAL视频信号（CVBS）、模拟R′G′B′或YPbPr视频信号。

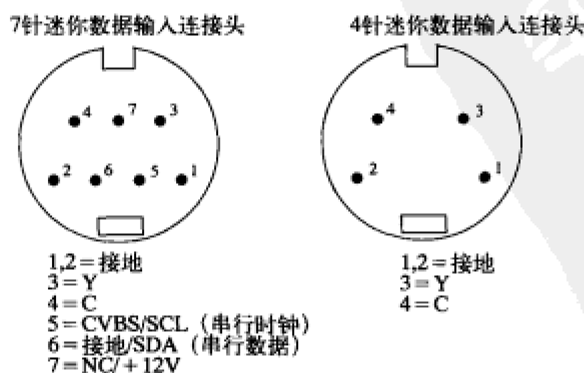


图5-1 S-Video连接头和信号命名

5.2 SCART接口

欧洲的大多数消费类视频分量都支持1到2个21针SCART连接头（通常也被称作Peritel、Peritelevision和Euroconnector）。这种连接准许模拟R′G′B′视频或S-Video、复合视频和模拟立体声音频在使用单根电缆的设备之间传输。复合视频信号必须一直存在，因为它需要为模拟R′G′B′视频信号提供基本的视频时序信息。注意，700mV的R′G′B′视频信号没有消隐基底电平和同步信息，如图5-4所示。

PAL VBI（场消隐期）数据可以按照576i复合视频信号的方式提供，将在第8章中对此进行讨论。

目前已经有几种类型的SCART引脚，取决于实现的具体功能，见表5-1~表5-3。图5-2给出了一些输出引脚的细节。

CENELEC EN 50049-1和IEC60933标准规定了基本的SCART连接头，包括信号电平。

69

表5-1 SCART连接头信号

引　脚	功　能	信号电平	阻　抗
1	右声道音频输出	0.5V rms	<1kΩ
2	右声道音频输入	0.5V rms	>10kΩ
3	左/单声道输出	0.5V rms	<1kΩ
4	接地，用于引脚1、2、3、6		
5	接地，用于引脚7		
6	左/单声道输入	0.5V rms	>10kΩ
7	蓝（或C）视频输入/输出	0.7V（或0.3V色同步电平）	75Ω
8	状态和宽高比输入/输出	9.5~12V＝4：3视频源 4.5~7V＝16：9视频源 0~2V＝无效视频源	>10kΩ
9	接地，用于引脚11		
10	数据2		
11	绿色视频输入/输出	0.7V	75Ω
12	数据1		
13	接地，用于引脚15		
14	接地，用于引脚16		
15	红（或C）视频输入/输出	0.7V（或0.3V色同步电平）	75Ω
16	RGB控制输入/输出	1~3V＝RGB 0~0.4V＝复合视频	75Ω
17	接地，用于引脚19		
18	接地，用于引脚20		
19	复合（或Y）视频输出	1V	75Ω
20	复合（或Y）视频输入	1V	75Ω
21	接地，用于引脚8、10、12、屏蔽		

注：通常，SCART1连接头支持复合视频和RGB，SCART2连接头支持复合视频和S-Video，SCART3连接头仅支持复合视频。SCART连接也可以用于将外部的解码器或解扰器加入视频处理路径中，视频信号从其中传出和回传到SCART连接头。

RGB控制信号控制TV在复合视频和RGB输入之间的切换，使得可以将图文电视（Teletext）叠加在视频上，甚至内部电视节目也可叠加在视频上。这使得外部字幕或图文电视广播的解码器能够将信息加入到当前的节目上。如果引脚16保持高电平，并给出标示RGB的信号，那么同步仍然是在复合视频的引脚上。某些设备（如DVD播放机）可以为SCART提供RGB，并保持引脚16恒为高电平。

当信源变为有效时，引脚8上将加载12V的电平。这使得电视自动地切换到SCART输入。当信源停止时，信号返回0V，并且恢复电视画面。如果提供16：9的变形节目，信源将引脚8上的信号电压提升到刚好6V。这导致电视机切换到SCART输入，并同时使得对变形的16：9节目进行视频处理。

70

表5-2　SDTV 10位R′G′B′颜色值

视频电平	消隐基底电平为7.5 IRE	消隐基底电平为0 IRE
白色	800	800
黑色	282	252
消隐	240	252
同步	16	16

表5-3　HDTV 10位R′G′B′颜色值

视频电平	消隐基底电平为0 IRE
白色	800
同步高	488
黑色	252
消隐	252
同步低	16

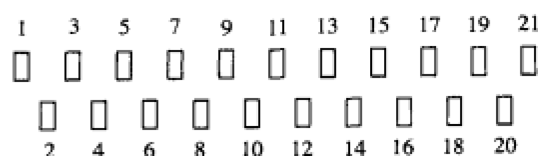


图5-2　SCART连接头

5.3　SDTV RGB接口

有些SDTV消费类视频设备支持模拟R′G′B′视频接口。NTSC和PAL VBI（垂直视频间隔）数据可以按照480i或576i R′G′B′视频信号提供，将在第8章对此进行讨论。使用的连接头为三个分离的RCA梅花连接头（消费类市场）或BNC连接头（专业视频和PC市场）。

水平或垂直视频时序取决于视频标准，在第4章已经对此进行了讨论。对于信号源，在连接头处视频信号的源阻抗应该为75Ω±5%。对于信号接收设备，视频输入应该是AC耦合的，并具有75Ω±5%的输入阻抗。三个信号在±5%ns内相互之间必须一致。

同步信息的提供方式如下：仅在绿色通道、所有三个通道、作为独立的复合同步信号或作为独立的水平和垂直同步信号。使用的gamma值为1/0.45。

5.3.1　7.5 IRE的消隐基底电平

如图5-3所示，标称有效视频幅度为714mV，包含一个7.5±2%IRE的消隐基底电平。206±6mV的复合同步信号可以仅在绿色通道中提供（消费类市场），或由所有三个通道提供（专业视频市场）。可以提供直到±1V的DC偏移。

1. 模拟R′G′B′数据生成

假定10位D/A转换器（DAC）的输出范围为0~1.305V（与第9章介绍的NTSC/PAL编码器使用的视频DAC相匹配），由10位的YCbCr到R′G′B′的转换公式为：

$$R' = 0.591\,(Y-64) + 0.810\,(Cr-512)$$

$$G' = 0.591\,(Y-64) - 0.413\,(Cr-512) - 0.199\,(Cb-512)$$

$$B' = 0.591\,(Y-64) + 1.025\,(Cb-512)$$

R′G′B′的标称10位颜色表示范围为0~518，这是为了与第9章介绍的NTSC/PAL编码器使用的有效视频电平相匹配。注意，在此应该支持R′G′B′的负值。

为了实现7.5 IRE的消隐基底电平，在有效视频期内，需要在数字R′G′B′数据上加上值42。而在消隐期内则加上0。

加上消隐基底电平后，R′G′B′数据被具有升余弦分布的消隐信号压缩了，这种压缩是为了降低视频信号开始和结束的摆率。对于480i和576i系统，消隐上升和下降的时间为140±20ns。对于480p和576p系统，消隐上升和下降的时间为70±10ns。

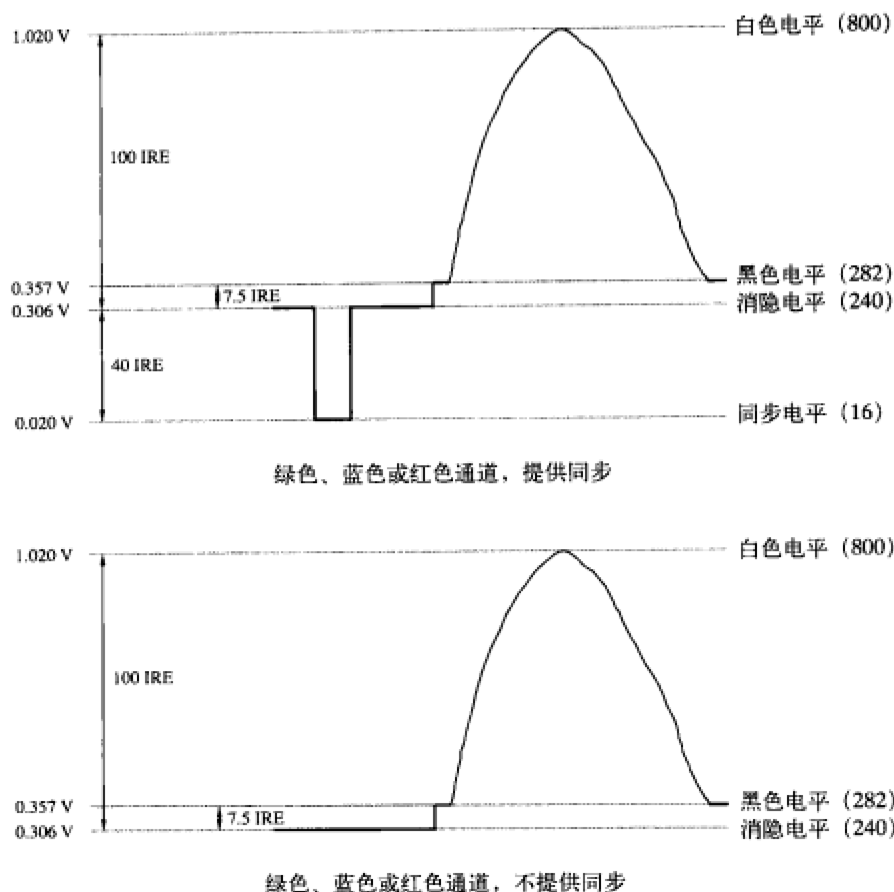


图5-3 SDTV模拟RGB电平。7.5 IRE的消隐电平

在消隐处理完成以后，复合同步信号可以加入到R′G′B′ 数据中。复合同步信号为16（提供同步）或240（没有同步）。应该对同步上升和下降时间进行处理以生成升余弦分布（在16与240之间），为此可以减缓同步信号的摆率。对于480i和576i系统，同步上升和下降的时间为140±20ns，水平同步在50%宽度点处的时间为4.7±0.1μs。对于480p和576p系统，同步上升和下降的时间为70±10ns，水平同步在50%宽度点处的时间为2.33±0.05μs。

到此为止，我们已经获得了带同步和消隐信息的数字R′G′B′ 数据了，如图5-3和表5-2所示。图5-3括号中的数字表示的是具有满量程输出值为1.305V的10位DAC的数据值。数字R′G′B′ 数据驱动三个10位DAC来生成模拟R′G′B′视频数据。 71

因为DAC的采样保持动作会引入$(\sin x)/x$特性，为了对此进行补偿，视频数据可以经过一个$[(\sin x)/x]^{-1}$滤波器进行数字滤波。另外，作为替代方法，由于在每个DAC之后，通常会提供一个模拟低通滤波器，因此矫正过程也可以在模拟滤波器中进行。

2. 模拟R′G′B′颜色数字化

假定10位A/D转换器（ADC）的输出范围为0~1.035V（为了与第9章介绍的NTSC/PAL编码器的视频DAC匹配），10位R′G′B′到YCbCr的转换公式为：

$$Y = 0.506(R'-282) + 0.992(G'-282) + 0.193(B'-282) + 64$$
$$Cb = -0.291(R'-282) - 0.573(G'-282) + 0.864(B'-282) + 512$$
$$Cr = 0.864(R'-282) - 0.724(G'-282) - 0.140(B'-282) + 512$$

R'G'B'具有标称的10位数，表示范围为282~800，这是为了与第9章介绍的NTSC/PAL解码器使用的有效视频电平相匹配。表5-2和图5-3给出了白色、黑色、消隐和（可选）同步信号电平的10位R'G'B'值。

5.3.2　0 IRE的消隐基底电平

如图5-4所示，标称有效视频信号幅度为700V，没有消隐基底电平。可以仅在绿色通道（消费类市场）或在所有三个颜色通道（专业视频市场）提供300±6mV的复合同步信号。可以提供直到±1V的DC偏移信号。

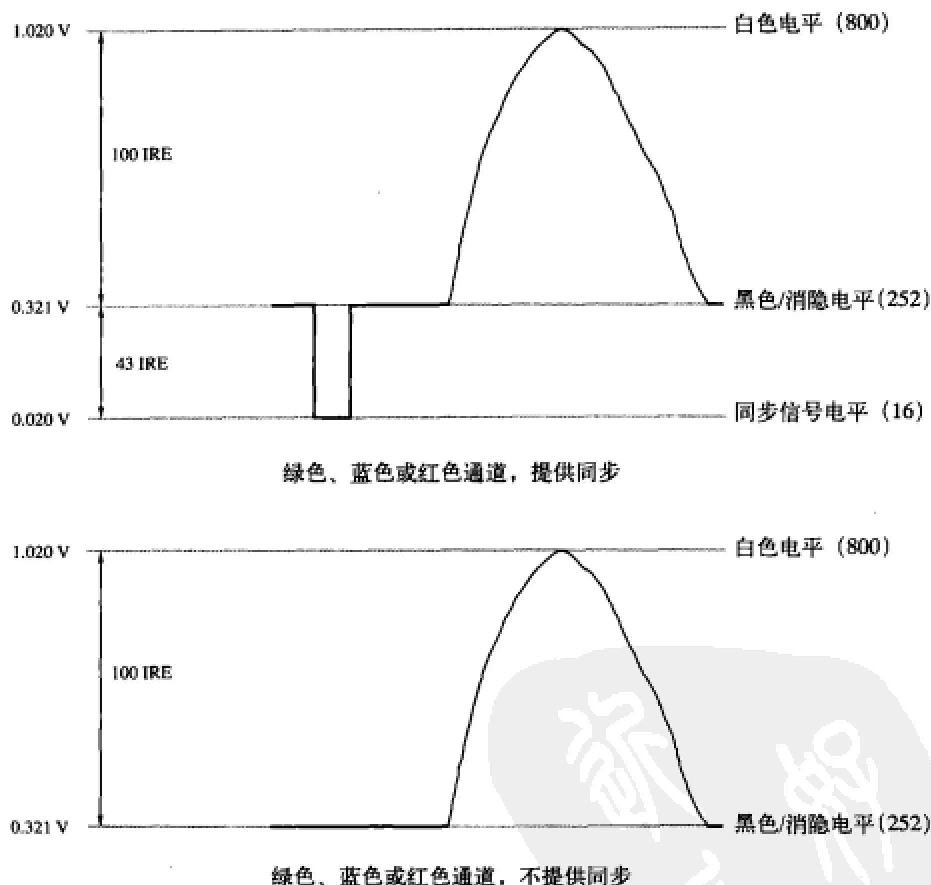


图5-4　SDTV模拟RGB电平，0 IRE的消隐电平

1. 模拟R'G'B'数据生成

假定10位的D/A转换器（DAC）具有输出范围为0~1.305V（与第9章介绍的NTSC/PAL编码器使用的视频DAC相匹配），由10位的YCbCr到R'G'B'的转换公式为：

$$R'=0.625\ (Y-64)\ +0.857\ (Cr-512)$$

$$G'=0.625\ (Y-64)\ -0.437\ (Cr-512)\ -0.210\ (Cb-512)$$

$$B'=0.625\ (Y-64)\ +1.084\ (Cb-512)$$

R'G'B'的标称10位颜色表示范围为0~548，这是为了与第9章介绍的NTSC/PAL编码器使用的有效视频电平相匹配。注意，在此应该支持负的R'G'B'值。

R'G'B'数据的处理与使用7.5 IRE消隐电平的情况一样。然而，在有效视频期内，没有增加消

隐电平，另外，同步值范围是16~252，而不是16~240。

到此为止，我们已经获得了带同步和消隐信息的数字R′G′B′数据了，如图5-4和表5-2所示。图5-4括号中的数字表示的是具有满量程输出值为1.305V的10位DAC的数据值。数字R′G′B′数据驱动三个10位DAC来生成模拟R′G′B′视频信号。

72
~
74

2. 模拟R′G′B′数字化

假定10位A/D转换器（ADC）的输出范围为0~1.035V（为了与第9章介绍的NTSC/PAL编码器的视频DAC匹配），10位R′G′B′到YCbCr的转换公式为：

$$Y=0.478\ (R'-252)\ +0.938\ (G'-252)\ +0.182\ (B'-252)\ +64$$

$$Cb=-0.275\ (R'-252)\ -0.542\ (G'-252)\ +0.817\ (B'-252)\ +512$$

$$Cr=0.817\ (R'-252)\ -0.685\ (G'-252)\ -0.132\ (B'-252)\ +512$$

R′G′B′具有标称的10位数，表示范围为252~800，这是为了与第9章介绍的NTSC/PAL解码器使用的有效视频电平相匹配。表5-2和图5-4给出了白色、黑色、消隐和（可选）同步信号电平的10位R′G′B′值。

5.4 HDTV RGB接口

有一些HDTV消费类视频设备支持模拟R′G′B′视频接口。HDTV RGB接口使用的是三个独立的RCA梅花连接头（消费类市场），或BNC连接头（专业视频或PC机市场）。

水平和垂直视频时序取决于特定的视频标准，如第4章讨论的一样。在源端，连接器内视频信号的源阻抗应该为75Ω±5%。在接收端，连接器内的视频输入应该是AC耦合的，并且输入阻抗为75Ω±5%。三个信号相互之间必须在±5ns内达到一致。

同步信号可以作为独立的复合同步信号，或作为独立的水平和垂直同步信号。同步信号可能仅出现在绿色通道上，也可能出现在所有的三个颜色通道上。使用的gamma值为1/0.45。

如图5-5所示，标称的有效视频信号的幅值为700mV，没有提供消隐电平。±300±6mV三极性电平可能仅出现在绿色通道上（消费类市场），也可能出现在所有3个颜色通道上（专业视频市场）。可以提供直到±1V的DC偏移。

5.4.1 模拟R′G′B′数据生成

假定10位DAC的输出范围为0~1.305V（为了与第9章介绍的NTSC/PAL编码器使用的视频DAC相匹配），10位YCbCr到R′G′B′的转换公式为：

$$R'=0.625\ (Y-64)\ +0.963\ (Cr-512)$$

$$G'=0.625\ (Y-64)\ -0.287\ (Cr-512)\ -0.114\ (Cb-512)$$

$$B'=0.625\ (Y-64)\ +1.136\ (Cb-512)$$

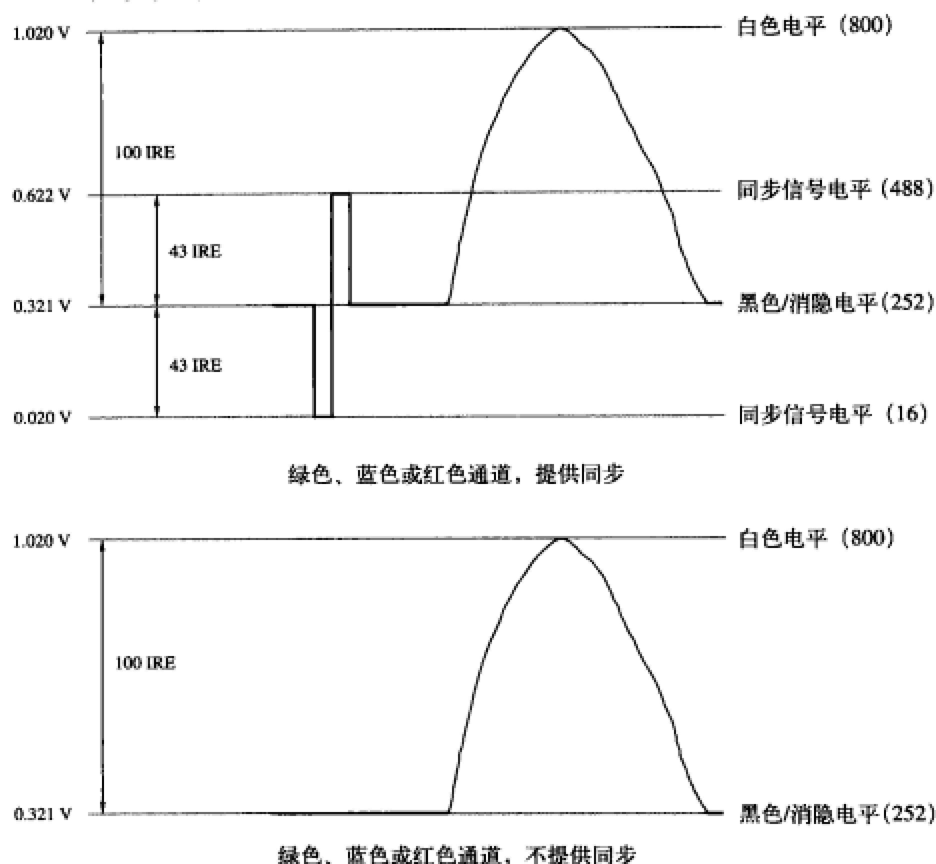
R′G′B′的标称10位颜色表示范围为0~548，这是为了与第9章介绍的NTSC/PAL编码器使用的有效视频电平相匹配。注意，在此应该支持R′G′B′的负值。

在加上消隐基底电平后，R′G′B′数据被具有升余弦分布的消隐信号压缩了，这种压缩是为了降低视频信号开始和结束的摆率。对于1080i和720p系统，消隐上升和下降的时间为54±20ns。对于1080p系统，消隐上升和下降的时间为27±10ns。

在消隐处理完成以后，复合同步信号可以加入到R′G′B′数据中。复合同步信号的值为16（同步低电平）、488（高电平同步）或252（没有同步）。应该对同步上升和下降时间进行处理以生成升余弦分布（在16与240之间），以减缓同步信号的摆率。对于1080i系统，同步上升和下降的时间为54±20ns，水平同步在50%宽度点处的时间为593±40ns。对于720p系统，同步上升和下降的时

75

间为54±20ns，水平同步在50%宽度点处的时间为539±40ns。对于1080p系统，同步上升和下降的时间为27±10ns，水平同步高和低在50%宽度点处的时间为296±20ns。



76

图5-5 HDTV模拟RGB电平，0 IRE的消隐电平

到此为止，我们已经获得了带同步和消隐信息的数字R′G′B′数据了，如图5-5和表5-3所示。图5-5括号中的数字表示的是具有满量程输出值为1.305V的10位DAC的数据值。数字R′G′B′数据驱动三个10位DAC来生成模拟R′G′B′视频数据。

5.4.2 模拟R′G′B′颜色数字化

假定10位的ADC具有的输入范围为0~1.305V（这是为了与第9章介绍的NTSC/PAL解码器所使用的视频ADC匹配），10位的R′G′B′转换为YCbCr的公式为：

$$Y = 0.341\ (R' - 252)\ + 1.143\ (G' - 252)\ + 0.115\ (B' - 252)\ + 64$$

$$Cb = -0.188\ (R' - 252)\ - 0.629\ (G' - 252)\ + 0.817\ (B' - 252)\ + 512$$

$$Cr = 0.817\ (R' - 252)\ - 0.743\ (G' - 252)\ - 0.074\ (B' - 252)\ + 512$$

R′G′B′具有标称的10位数，表示范围为252~800，这是为了与第9章介绍的NTSC/PAL解码器使用的有效视频电平相匹配。表5-3和图5-5展示了白色、黑色、消隐和（可选）同步信号电平的10位R′G′B′值。

5.4.3 约束图像

由于使用高清晰模拟接口的复制保护技术受到一定的限制，因此，有一些标准和DRM的实

现仅准许输出约束图像（constrained image）。尽管视频采样数和视频时序的总数维持不变（例如，1280×720p或1920×1080i），但约束图像的最大有效分辨率只为960×540p。

在这种情况下，通过一个授权的安全数字视频输出（如HDMI），完全分辨率的图像仍然是可以获得的。

5.5 SDTV YPbPr接口

一些SDTV消费类视频设备支持模拟YPbPr视频接口。NTSC和PALVBI数据可以在480i或576i的Y视频信号上提供，这在第8章讨论。SDTV YPbPr接口采用三个分离的RCA梅花连接头（消费类市场）或BNC连接头（专业视频市场）。

水平和垂直视频时序取决于视频标准，这已经在第4章讨论过了。对于源信号，在连接头处视频信号的源阻抗为75Ω±5%。对于接收设备，视频输入应该是AC耦合的，并具有75Ω±5%的输入阻抗。三个信号相互之间必须在±5ns以内达到一致。 77

对于消费类产品，仅在Y通道上提供复合同步。对于专业视频应用，在三个通道上都提供复合同步信号。指定的校准参数gamma为1/0.45。

如图5-6和图5-7所示，Y信号由700mV的有效视频（没有消隐基底电平）组成。Pb和Pr峰到峰的幅度为700mV。3006mV的复合同步信号在Y通道（消费市场）和所有三个通道（专业视频市场）提供。可以提供直到±1V的DC偏移。表5-4和表5-5给出了100%和75%的YPbPr彩条值。

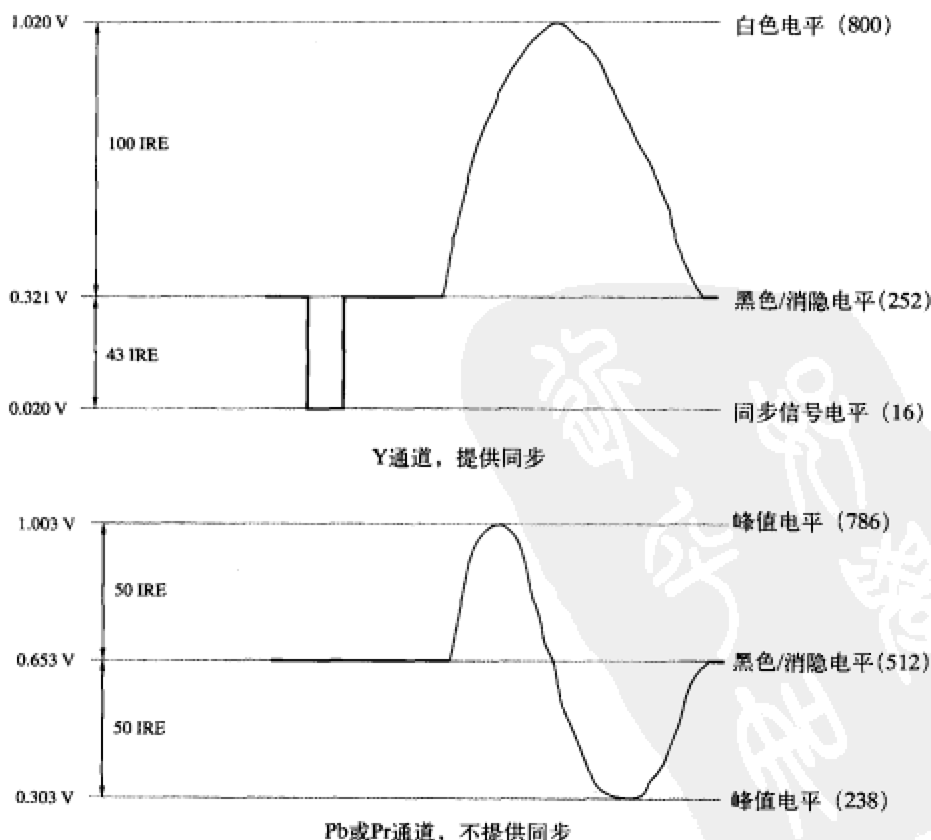
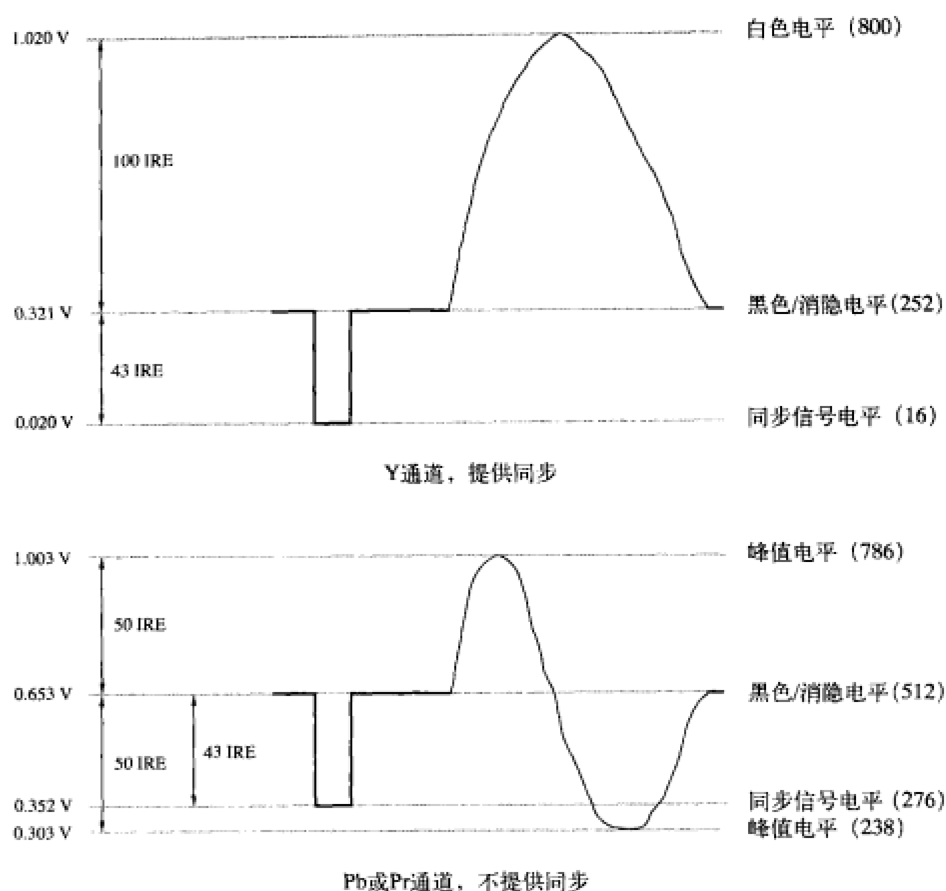


图5-6 EIA-770.2 SDTV模拟YPbPr电平，同步信号在Y上 78



79

图5-7 SDTV模拟YPbPr电平，同步信号在YPbPr上

表5-4 EIA-770.2 SDTV YPbPr和YCbCr 100%彩条。YPbPr相对消隐电平的值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	100	88.6	70.1	58.7	41.3	29.9	11.4	0
	mV	700	620	491	411	289	209	80	0
Pb	IRE	0	−50	16.8	−33.1	33.1	−16.8	50	0
	mV	0	−350	118	−232	232	−118	350	0
Pr	IRE	0	8.1	−50	−41.8	41.8	50	−8.1	0
	mV	0	57	−350	−293	293	350	−57	0
Y	64~940	940	840	678	578	426	326	164	64
Cb	64~960	512	64	663	215	809	361	960	512
Cr	64~960	512	585	64	137	887	960	439	512

表5-5 EIA-770.2 SDTV YPbPr和YCbCr 75%彩条。YPbPr相对消隐电平的值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	75	66.5	52.6	44	31	22.4	8.6	0
	mV	525	465	368	308	217	157	60	0
Pb	IRE	0	−37.5	12.6	−24.9	24.9	−12.6	37.5	0
	mV	0	−262	88	−174	174	−88	262	0

（续）

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Pr	IRE	0	6.1	−37.5	−31.4	31.4	37.5	−6.1	0
	mV	0	43	−262	−220	220	262	−43	0
Y	64~940	721	646	525	450	335	260	139	64
Cb	64~960	512	176	625	289	735	399	848	512
Cr	64~960	512	567	176	231	793	848	457	512

80

5.5.1 模拟YPbPr数据生成

假定10位DAC的输出范围为0~1.305V（为了与第9章介绍的NTSC/PAL编码器使用的视频DAC相匹配），10位YCbCr到YPbPr的转换公式为：

Y＝［（800－252）/（940－64）］（Y－64）

Pb＝［（800－252）/（960－64）］（Cb－512）

Pr＝［（800－252）/（960－64）］（Cr－512）

Y具有10位，其标称范围为0~548，这是为了与第9章要介绍的NTSC/PAL编码器使用的有效视频电平相匹配。Pb和Pr有10位，标称范围为0~±274。注意，在此应该支持Y的负值。

具有升余弦分布的消隐信号压缩了YPbPr数据，这种压缩是为了减缓视频信号开始和结束的摆率。对于480i和576i系统，消隐上升和下降的时间为140±20ns。对于480p和576p系统，消隐上升和下降的时间为70±10ns。

在消隐处理完成以后，在Y数据中加入了复合同步信号。复合同步信号的值为16（提供同步）或252（没有同步）。应该对同步上升和下降时间进行处理以生成升余弦分布（在276与512之间），来减缓同步信号的摆率。

在消隐处理完成以后，也可以将复合同步信息加入PbPr数据。复合同步信号的值为16（提供同步）或252（没有同步）。应该对同步上升和下降时间进行处理以生成升余弦分布（在276与512之间），来减缓同步信号的摆率。

对于480i和576i系统，同步升降时间为140±20ns，水平同步在50%宽度点处的时间为4.7±0.1μs。对于576p系统，同步上升和下降的时间为7010ns，水平同步在50%宽度点处的时间为2.33±0.05μs。

到此为止，我们已经获得了带同步和消隐信息的数字YPbPr数据了，如图5-6、图5-7和表5-6所示。图5-6和图5-7括号中的数字表示的是具有满量程输出值为1.305V的10位DAC的数据值。数字YPbPr数据驱动三个10位DAC来生成模拟YPbPr视频数据。

表5-6 SDTV 10位YPbPr值

视频电平	Y	PbPr
白色	800	512
黑色	252	512
消隐	252	512
同步	16	276

81

5.5.2 模拟YPbPr数字化

假定10位ADC的输入范围为0~1.305V（为了与第9章将要介绍的NTSC/PAL解码器使用的视频ADC匹配），10位YPbPr到YCbCr的转换公式为：

$$Y = 1.5985\ (Y - 252) + 64$$
$$Cb = 1.635\ (Pb - 512) + 512$$
$$Cr = 1.635\ (Pr - 512) + 512$$

10位Y值的标称表示范围为252~800，这是为了与第9章将要介绍的NTSC/PAL解码器使用的有效视频电平相匹配。表5-6、图5-6和图5-7展示了白色、黑色、消隐和（可选）同步信号电平的10位YPbPr值。

5.5.3　480p系统的VBI数据

1. CGMS A类保护

CEA-805、IEC 61880-2和EIA-J CPR-1024-1定义了CGMS（复制代次管理系统）数据的格式，该定义位于480p系统的第41行。其波形如图5-8所示。

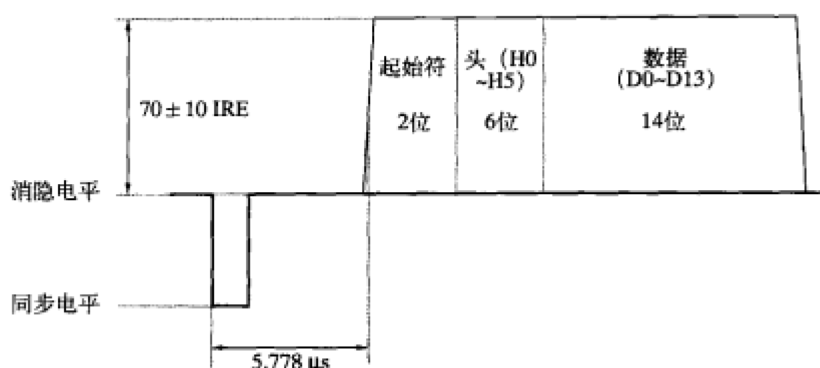


图5-8　CEA-805、IEC 61880-2和EIA-J CPR-1204-1线41时序

使用的是27MHz（59.94Hz帧率）或27.027MHz（60Hz的帧率）的采样时钟频率。每个数据为26个时钟周期，或963±30ns，最大的上升和下降时间跨度为50ns。逻辑1的幅度为70±10 IRE，逻辑0的幅度为0±5 IRE。

在0_H后，2位的起始符开始于156时钟周期或5.778μs，起始符为10。

6位的头紧随起始符，头定义了载荷数据的属性，如表5-7所示。在任何数据服务的最后一包之后紧接着信息结束符，数据服务需要占据2个以上的数据包。结束符是连续的数据位，其值皆为0。ECCI为数据服务，需要占据2个以上的数据包，因此需要使用消息结束符。

表5-7　CEA-805、IEC 61880-2和EIA-J CPR-1204-1线41头格式。如果标示为A类包，H2~H5数据位必须为“0000”

H0	H1	宽　高　比	图像显示格式
0	0	4：3	普通
0	1	4：3	Letterbox
1	0	16：9	普通
1	1	CEA-805 A类包	

H2	H3	H4	H5	服　务　名
0	0	0	0	CGMS（见表5-8）
0	0	0	1	扩展复制控制信息（ECCI）
0	0	1	0	保留
	⋮			
1	1	1	0	
1	1	1	1	信息结束符（如果没有版权信息，为默认设置）

82
~
83

CGMS数据占14位载荷，具体见表5-8。14位ECCI数据载荷当前被保留，由全1组成。

表5-8 CEA-805、IEC 61880-2和EIA-J CPR-1204-1线41CGMS服务格式

D0	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13
G0	G1	G2	G3	ASB	0	0	0	CRC $=x^6+x+1$					

G0~G1：CGMS定义

- 00 准许复制
- 01 不能再复制（已经制作了一份副本）
- 10 准许复制一次
- 11 不准复制

G2~G3：模拟保护服务（仅当G0~G1为“01”或“10”时有效）

- 00 没有模拟保护服务
- 01 PSP开，脱落的颜色
- 10 PSP开，2线颜色脱落
- 11 PSP开，4线颜色脱落

ASB：模拟源位

- 0 没有模拟预录制媒体
- 1 模拟预录制媒体

2. CGMS B类保护

CEA-805在408p系统的40线定义了CGMS（复制代次管理系统）数据的格式。其波形见图5-9。

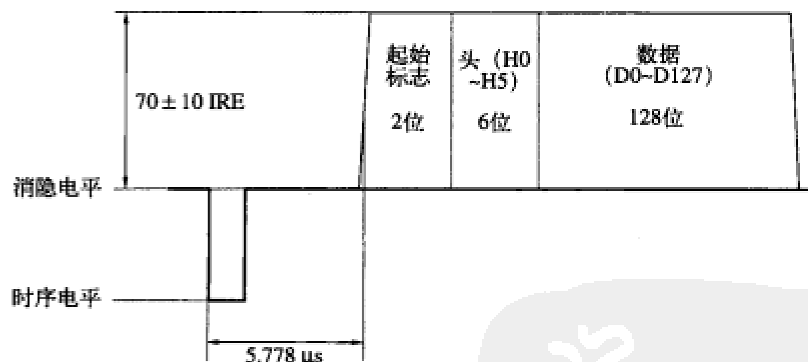


图5-9 CEA-805线40时序

采样时钟频率为27MHz（59.94Hz帧率）或27.027MHz（60Hz帧率）。每个数据位占4个时钟周期，或148±18.5ns，最大上升和下降时间跨度为37ns。逻辑“1”的幅度为70±10IRE，逻辑“0”的幅度为0±5IRE。

在0_H后，2位的起始符开始于156时钟周期或5.778μs，起始符为“10”。

6位的头紧随起始符，头定义了载荷数据的属性，如表5-9所示。

表5-9 CEA-805 线40的头格式

H0	H1	H2	H3	H4	H5	服务名
0	0	0	0	0	0	保留给以后使用
		⋮				
1	1	0	0	0	1	
1	1	0	0	1	0	B类包
1	1	0	0	1	0	保留给以后使用
		⋮				
1	1	1	1	1	1	

84

16字节的载荷如表5-10所示。

表5-10 CEA-805 线40载荷格式

D7	D6	D5	D4	D3	D2	D1	D0
版本号＝0000 0001							
载荷包长度＝0001 0000							
AR1	AR0	ASB	A0	1	B0	S1	S0
C3	C2	C1	C0	R3	R2	R1	R0
RCI	1	1	1	G3	G2	G1	G0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
顶部条结束的线编码（低8位）							
顶部条结束的线编码（高8位）							
底部条开始的线数编码（低8位）							
底部条开始的线数（高8位）							
左部条结束的像素数（低8位）							
左部条结束的像素数（高8位）							
右部条结束的像素数（低8位）							
右部条结束的像素数（高8位）							
1	1	CRC＝x^6+x+1					

AR1~AR0：显示器宽高比的设计值
- 00　4：3 普通
- 01　4：3 Letterbox
- 10　16：9 普通
- 11　保留

ASB：模拟源位

A0：有效格式描述数据（AFD）标志
- 0　没有AFD数据（R0~R3）
- 1　AFD数据（R0~R3）有效

B0：条数据（用于Letterbox）
- 0　没有条数据（bar data）
- 1　提供条数据

S1~S0：扫描数据（过扫描和欠扫描数目不显示）
- 00　没有数据
- 01　过扫描（电视）
- 10　欠扫描（计算机）
- 11　保留

C3~C0：比色法
- 0000　无数据
- 0001　BT.601
- 0010　BT.709
- 0011　保留
- ⋮　⋮
- 1111　保留

R0-R3：有效格式描述（AFD）有效格式数据（参考表13-56）

RCI：重新发布控制信息（RCI）标志

G0~G1：CGMS定义（参考表5-8）

G2~G3：模拟保护服务（参考表5-8）

85～86

5.5.4 576p系统的VBI数据

1. CGMS

IEC 62375为576p系统在线43上定义了CGMS（复制代次管理系统）和宽屏信号（WSS）数据格式，其波形见图5-10。该标准准许WSS增强的16：9 TV以正确的宽高比显示节目。

2. 数据时序

CGMS和WSS数据通常在43线上，如图5-10所示。但是，由于视频编辑的缘故，数据可以位于43~47之间的任何线上。

时钟频率为10MHz（±1kHz）。信号波应该是正弦方波脉冲，半幅度持续的时间为100±10ns。信号幅度为500mV±5%。

NRZ数据位是通过双相位编码调制器进行处理的，这使得在10MHz时一个数据周期等于10个数据位。

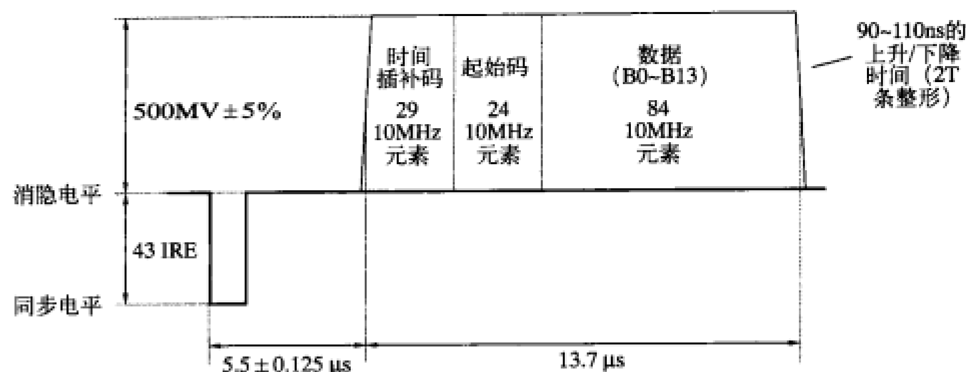


图5-10 IEC 62375线43 CGMS时序

3. 数据内容

WSS由时间插补码、起始码和14位的数据组成，见表5-11。

87

表5-11 IEC 62375 43线上的WSS信息

时间插补码	10MHz的29个元素	1 1111 0001 1100 0111 0001 1100 0111 (0x1F1C71C7)
起始码	10MHz的24个元素	0001 1110 0011 1100 0001 1111 (0x1E3C1F)
组1数据（宽高比）	10MHz的24个元素 “0” =000 111 “1” =111 000	b0、b1、b2、b3
组2数据（增强服务）	10MHz的24个元素 “0” =000 111 “1” =111 000	b4、b5、b6、b7 （由于b4、b5、b6和b7是保留的，故都等于0）
组3数据（字幕）	10MHz的18个元素 “0” =000 111 “1” =111 000	b8、b9、b10 （由于是保留的，故b8＝“0”）
组4数据（保留）	10MHz的18个元素 “0” =000 111 “1” =111 000	b11、b12、b13

❑ 时间插补码

时间插补（run-in）码由29个10MHz的特定序列的元素组成，见表5-11。

❑ 起始码

起始码（start code）由24个10MHz的特定序列的元素组成，见表5-11。

❑ 组1数据

组1数据（group 1 data）由4个数据位组成，其内容指定了宽高比。每个数据位生成10MHz的6个原色。b0为LSB（最低有效位）。

表5-11列出了数据位的指派和使用。表5-12中列出了特定宽高比（a=1.33、1.56、1.78）的有效线数目。

表5-12 IEC 62375 第1组（宽高比）数据位的指派和使用

b0、b1、b2、b3	宽高比标注	格 式	在4：3显示器上的位置	有效线	最低要求
0001	4：3	全屏格式	–	576	情形1

（续）

b0、b1、b2、b3	宽高比标注	格　式	在4：3显示器上的位置	有效线	最低要求
1000	14：9	Letterbox模式	居中	504	情形2
0100	14：9	Letterbox模式	靠近顶部	504	情形2
1101	16：9	Letterbox模式	居中	430	情形3
0010	16：9	Letterbox模式	靠近顶部	430	情形3
1011	>16：9	Letterbox模式	居中	–	情形4
0111	14：9	全屏格式	居中	576	–
1110	16：9	全屏格式（变形）	–	576	–

宽高比标注表示的是可能的宽高比（a）范围和有效线数目：

4：3	$a \leq 1.46$	527~576
14：9	$1.46 < a \leq 1.66$	463~526
16：9	$1.66 < a \leq 1.90$	405~462
>16：9	$a > 1.90$	<405

为了准许自动选择模式，16：9的接收设备应该支持下面的最低要求。

情形1：4：3宽高比的画面应该定位于显示器的正中央，让黑色带位于显示器的左边和右边。

情形2：14：9宽高比的画面应该定位于显示器的正中央，让黑色带位于显示器的左边和右边。替代方法是，图像可以以全屏方式进行显示，准许小的（典型情况下为8%）水平几何误差。

情形3：16：9宽高比的画面应该定位于显示器的正中央，让黑色带位于显示器的左边和右边。

情形4：大于16：9宽高比的画面应该以情形1进行显示，或通过在垂直方向上放大以显示器的高度来显示。

❑ 组3数据

组3数据（group 3 data）由3个数据位组成，这些位规定了字幕属性。每个数据位以10MHz的频率生成了6个元素。数据位b8为LSB（最低有效位）。

b9、b10：开放字幕。00：没有。01：在有效画面之外。10：在有效画面内。11：保留。

❑ 组4数据

组4数据（group 4 data）由3个数据位组成，它们规定了环绕声和复制保护。每个数据位以10MHz的频率生成6个元素。数据位b11为LSB（最低有效位）。

b11：环绕声。0：没有。1：有。

b12：版权。0：没有版权声明或不知道。1：版权声明。

b13：复制保护。0：复制不受限制。1：复制受限。

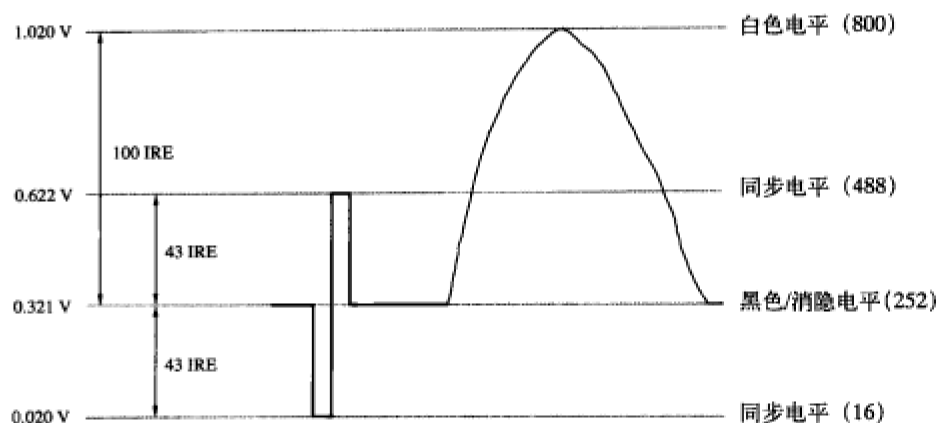
5.6　HDTV YPbPr接口

大多数HDTV消费类视频设备支持一个模拟YPbPr视频接口。该接口使用的是三个独立的RCA梅花连接头（消费类市场）或BNC连接头（专业视频市场）。

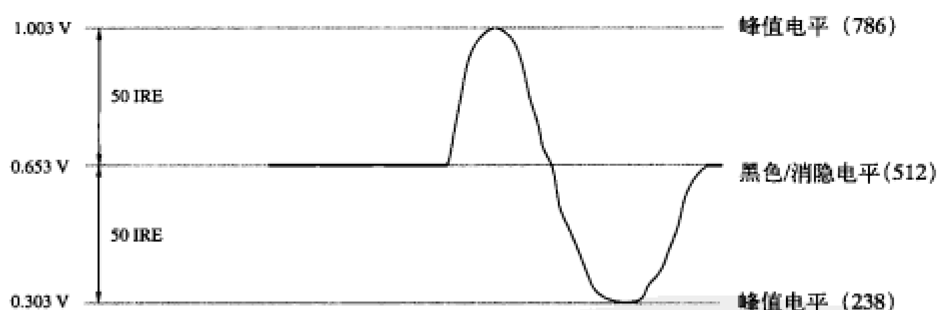
水平和垂直视频时序取决于视频标准，如第4章讨论的一样。对于源信号，视频信号在连接头处应该具有源阻抗75Ω±5%。对于接收设备，视频输入应该是AC耦合的，具有的输入阻抗为75Ω±5%。三个信号之间的协同误差要在5ns内。

对于消费类产品，仅在Y通道提供复合同步信号。对于专业视频应用，同时在三个通道提供复合同步。指定的校正gamma值为1/0.45。

如图5-11和图5-12所示，Y信号由700mV的有效视频（没有消隐基底电平）组成。Pb和Pr峰到峰的幅度为700mV。±300±6mV的复合同步信号仅在Y通道提供（消费市场）或在所有三个通道提供（专业视频市场）。可以提供直到±1V的DC偏移。100%和75%YPbPr彩条的值见表5-13和表5-14。



Y通道，同步显示



Pb或Pr通道，不同步显示

图5-11 EIA-770.3 HDTV模拟YPbPr电平，同步信号在Y通道

表5-13 EIA-770.3 HDTV YPbPr和YCbCr 100%彩条。YPbPr的值是相对消隐电平的

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	100	92.8	78.7	71.5	28.5	21.3	7.2	0
	mV	700	650	551	501	200	149	50	0
Pb	IRE	0	−50	11.4	−38.5	38.5	−11.4	50	0
	mV	0	−350	80	−270	270	−80	350	0
Pr	IRE	0	4.6	−50	−45.4	45.4	50	−4.6	0
	mV	0	32	−350	−318	318	350	−32	0
Y	64~940	940	877	753	690	314	251	127	64
Cb	64~960	512	64	614	167	857	410	960	512
Cr	64~960	512	553	64	106	918	960	471	512

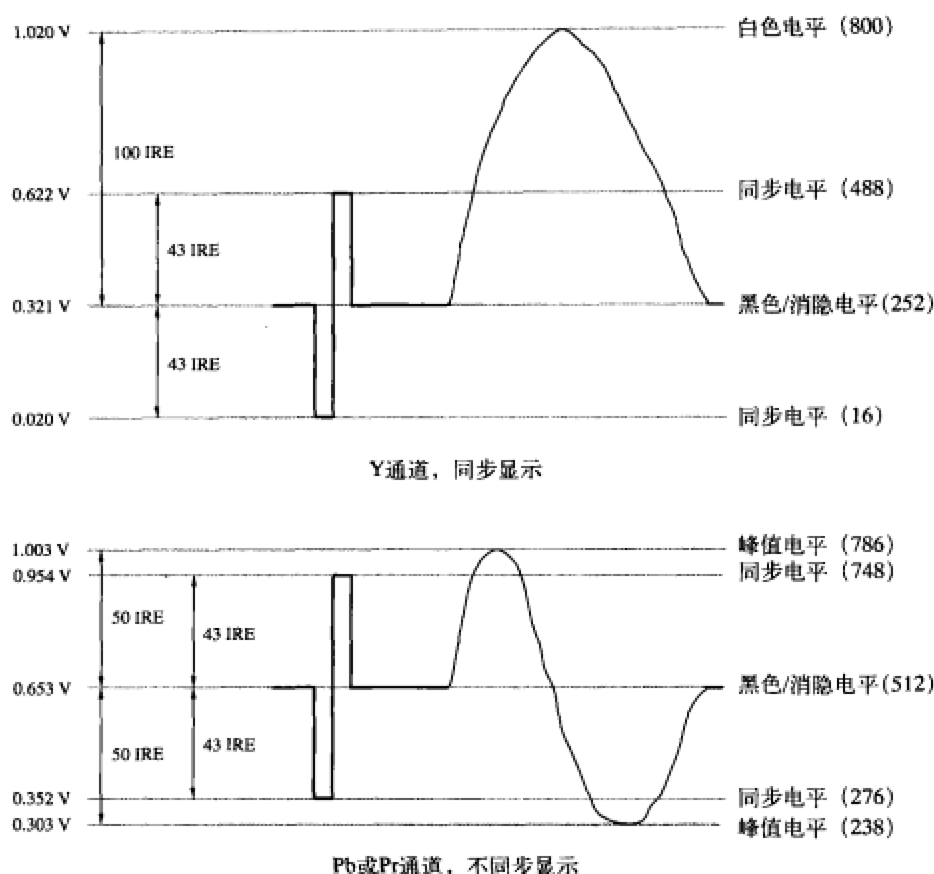


图5-12 SMPTE 274M和296M HDTV模拟YPbPr电平，在PbPr上同步

表5-14 EIA-770.3 HDTV YPbPr和YCbCr 75%彩条。YPbPr的值是相对消隐电平的

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	75	69.6	59	53.7	21.3	16	5.4	0
	mV	525	487	413	376	149	112	38	0
Pb	IRE	0	−37.5	8.6	−28.9	28.9	−8.6	37.5	0
	mV	0	−263	60	−202	202	−60	263	0
Pr	IRE	0	3.5	−37.5	−34	34	37.5	−3.5	0
	mV	0	24	−263	−238	238	263	−24	0
Y	64~940	721	674	581	534	251	204	111	64
Cb	64~960	512	176	589	253	771	435	848	512
Cr	64~960	512	543	176	207	817	848	481	512

5.6.1 模拟YPbPr数据生成

假定10位DAC的输出范围为0~1.305V（这是为了与第9章要介绍的NTSC/PAL使用的视频DAC相匹配），10YCbCr到YPbPr的转换公式为：

$$Y = [(800-252)/(940-64)](Y-64)$$

$$Pb = [(800-252)/(960-64)](Cb-512)$$

Pr=[（800−252）/（960−64）]（Cr−512）

Y值具有10位，其标称表示范围为0~548，这是为了与第9章要介绍的NTSC/PAL编码器使用的有效视频电平相匹配。Pb和Pr具有10位数据，标称表示范围为0~±274。注意，在此，应该支持Y为负值。

消隐信号压缩了YPbPr数据，消隐信号具有升余弦分布，这是为了减缓视频信号开始与结束的摆率。对于1080i和720p系统，消隐上升和下降时间为54±20ns。对于1080p系统，消隐上升和下降的时间为27±10ns。

复合同步信息在消隐处理完成以后可以加到Y数据上。指定同步信号的值为16（同步低）、488（同步高）或252（没有同步）。应该对同步上升和下降的时间进行处理，以生成升余弦分布来减缓同步信号的摆率。

复合同步信息在消隐处理完成以后可以加到PbPr数据上。指定同步信号的值为276（同步低）、748（同步高）或512（没有同步）。应该对同步上升和下降的时间进行处理，以生成升余弦分布来减缓同步信号的摆率。

对于1080i系统，同步上升和下降时间为54±20ns，水平同步在50%幅度的高低宽度为539±40ns。对于1080p系统，同步上升和下降时间为27±10ns，水平同步在50%幅度的高低宽度为296±20ns。

至此，我们已经具有了带同步和消隐信息的数字YPbPr，如图5-11、图5-12以及表5-15所示。图5-11和图5-12中括号中的数字表示的是具有满量程输出值（full-scale output）为1.305V的10位DAC的数据值。数字YPbPr数据驱动三个10位DAC生成模拟YPbPr视频信号。

89~90

表5-15 DHTV 10位YPbPr值

视频电平	Y	PbPr
白色	800	512
同步高	488	748
黑色	252	512
消隐	252	512
同步低	16	276

5.6.2 模拟YPbPr数字化

假定10位ADC的输出范围为0~1.305V（为了与第9章要介绍的NTSC/PAL使用的视频ADC相匹配），10位YPbPr到YCbCr的转换公式为：

Y=1.5985（Y−252）+64

Cb=1.635（Pb−512）+512

Cr=1.635（Pr−512）+512

Y具有10个数据位，标称表示范围为252~800，这是为了与第9章要介绍的NTSC/PAL解码器使用的有效视频电平相匹配。表5-15、图5-11和图5-12表示的是白色、黑色、消隐和（可选）同步电平的10位YPbPr值。

5.6.3 720p系统的VBI数据

1. CGMS A类

对于720p系统，CEA-805和EIA-J CPR-1204-2在线24上定义了CGMS（复制代次管理系统）数据格式。其波形见图5-13。

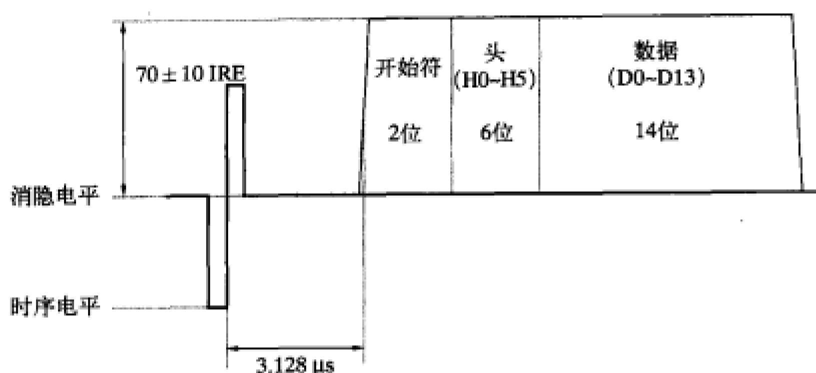


图5-13　CEA-805和EIA-J CPR-1204-2线24时序

采样时钟频率使用的是74.176MHz（59.94Hz帧率）或74.25MHz（60Hz帧率）。每个数据位为58个时钟周期（782±30ns），宽为最大上升和下降时间50ns。逻辑“1”的幅度为70±10 IRE；逻辑“0”的幅度为0±5 IRE。

在0_H后，2位起始符开始后占了232周期，或大约3.128μs。起始符的内容为“10”。

6位的头和14位的CGMS载荷数据格式与本章前面讨论过的480p系统是一样的。

2. CGMS B类

对于720p系统，CEA-805在线234上定义了CGMS（复制代次管理系统）数据格式。其波形见图5-14。

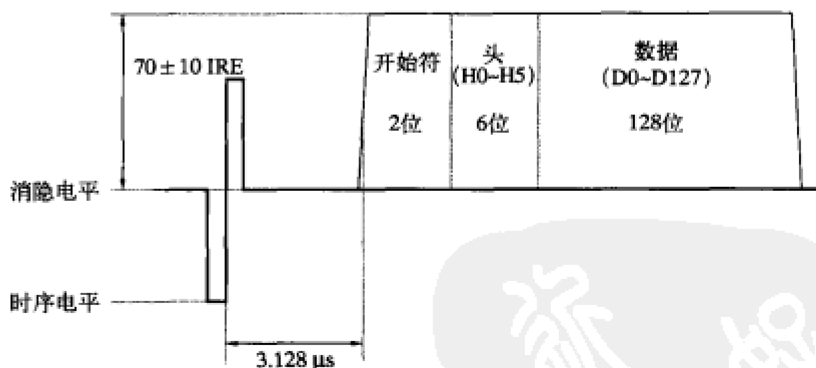


图5-14　CEA-805线23时序

采样时钟频率使用的是74.176MHz（59.94Hz帧率）或74.25MHz（60Hz帧率）。每个数据位为8个时钟周期（107.7±18.5ns），宽为最大上升和下降时间37ns。逻辑“1”的幅度为70±10 IRE；逻辑“0”的幅度为0±5 IRE。

91 ~ 94

在0_H后，2位起始符开始后占了232周期，或大约3.128μs。起始符的内容为“10”。

6位的头和16字节的CGMS载荷数据格式与本章前面讨论过的480p系统是一样的。

5.6.4　1080i系统的VBI数据

1. CGMS A类

对于1080i系统，CEA-805和EIA-J CPR-1204-2在线19和582上定义了CGMS（复制代次管理系统）数据格式。其波形见图5-15。

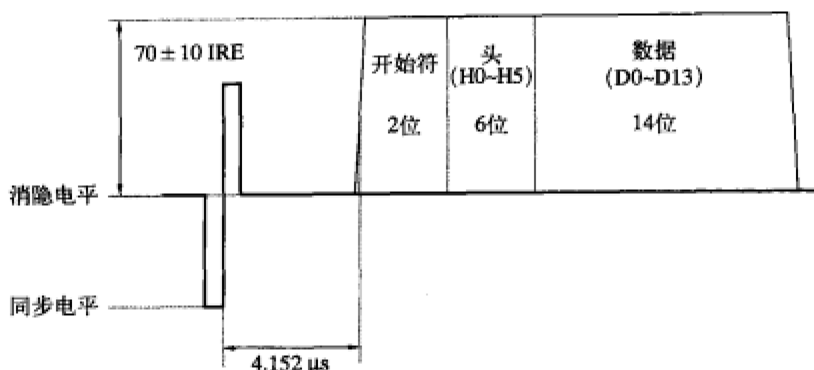


图5-15 CEA-805和EIA-J CPR-1204-2线19和582时序

采样时钟频率使用的是74.176MHz（59.94Hz帧率）或74.25MHz（60Hz帧率）。每个数据位为77个时钟周期（1038±30ns），宽为最大上升和下降时间50ns。逻辑"1"的幅度为70±10 IRE；逻辑"0"的幅度为0±5 IRE。

在0_H后，2位起始符开始后占了380周期，或大约4.152μs。起始符的内容为"10"。

6位的头和14位的CGMS载荷数据格式与本章前面讨论过的480p系统是一样的。

2. CGMS B类

对于1080i系统，CEA-805在线18和581上定义了CGMS（复制代次管理系统）数据格式。其波形见图5-16。

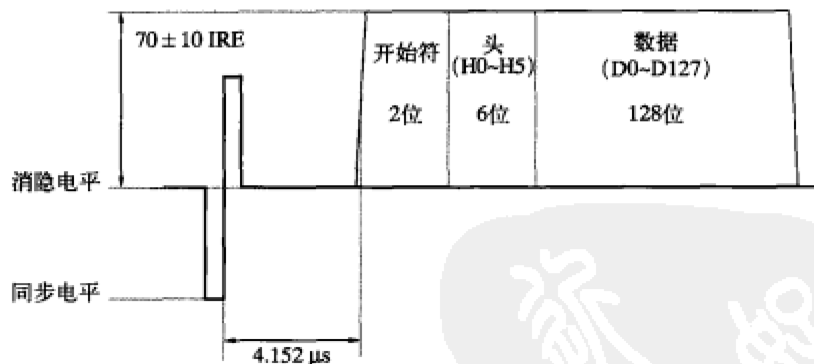


图5-16 CEA-805线18和581时序

采样时钟频率使用的是74.176MHz（59.94Hz帧率）或74.25MHz（60Hz帧率）。每个数据位为10个时钟周期（134.6±18.5ns），宽为最大上升和下降时间37ns。逻辑"1"的幅度为70±10 IRE，逻辑"0"的幅度为0±5 IRE。

在0_H后，2位起始符开始后占了308周期，或大约4.152μs。起始符的内容为"10"。

6位的头和16字节的CGMS载荷数据格式与本章前面讨论过的480p系统是一样的。

5.6.5 约束图像

由于针对高清晰模拟接口的复制保护技术受到一定限制，因此，有些标准和DRM实现仅准许输出约束图像（constrainted image）。尽管总的视频采样数目和视频时序并没有改变（例如，1280×720p或1920×1080i30），但约束图像有效的最大分辨率只能达到860×540p。

在此情况下，通过授权的安全数字视频输出（如HDMI），仍然可以获取满量程分辨率图像。

5.7 D型连接头接口

在日本和新加坡，在某些高端消费类设备上，14针母D型连接头（EIA-J CP-4120标准、EIA-J RC-5237连接头）是选择性可用的。主要是用于转换EIA 770.2或EIA770.3隔行或逐行模拟YPbPr视频数据。

D型连接头有5种类型，分别称为D1、D2、D3、D4和D5，每一种表示所支持的视频格式，如表5-16所示。图5-17表示的是该连接头，表5-17列出了引脚的名称。

表5-16 D型连接头支持的视频格式

	480i	480p	720p	1080i	1080p
D1	×				
D2	×	×			
D3	×	×		×	
D4	×	×	×	×	
D5	×	×	×	×	×

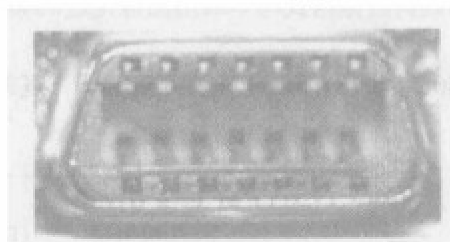


图5-17 D型连接头

表5-17 D型连接头引脚描述

引 脚	功 能	信号电平	阻 抗
1	Y	0.700V＋同步	75Ω
2	地－Y		
3	Pb	±0.350V	75Ω
4	地－Pb		
5	Pr	±0.350V	75Ω
6	地－Pr		
7	保留1		
8	线1	0V、2.2V或5V[1]	10k±3kΩ
9	线2	0V、2.2V或5V[1]	10k±3kΩ
10	保留2		
11	线3	0V、2.2V或5V[1]	10k±3kΩ
12	地－检测是否插入		
13	保留3		
14	检测是否插入	0V表示已经插入[2]	>100kΩ

1. 2.2V的范围为2.2V±0.8V，5V的范围为5V±1.5V。
2. 在设备内部，引脚12是接地的，引脚14是通过一个电阻抬高到5V的。在每个D型连接头的插头内部，引脚12和14都是短路在一起的。

三线信号（线1、线2和线3）表示的是YPbPr源视频的分辨率和帧率，如图5-18所示。

97

表5-18 D型连接头的各种视频格式线信号电压电平

分辨率		帧率	线1 扫描线	线2 帧率	线3 宽高比	色度和参考白色	颜色空间变换公式	Gamma校准	Y上的同步信号幅度
1920×1080		30i	5V	0V	5V	EIA-770.3	EIA-770.3	EIA-770.3	±0.300 V[3]
		25i[2]	5V	2.2V	5V				
		30p	5V	2.2V	5V				
		25p[2]	5V	2.2V	5V				
		24p[2]	5V	2.2V	5V				
		24sF[2]	5V	2.2V	5V				
1280×720		60p	2.2V	5V	5V				
		50p[2]	2.2V	2.2V	5V				
		30p	2.2V	2.2V	5V				
		25p[2]	2.2V	2.2V	5V				
		24p[2]	2.2V	2.2V	5V				
640×480		60p[2]	0V	5V	0V	EIA-770.2	EIA-770.2	EIA-770.2	−0.300 V[3]
720×480	16：9 Squeeze	60p	0V	5V	5V				
	16：9 Squeeze	30i	0V	0V	5V				
	16：9 Letterbox	30i	0V	0V	2.2V				
	4：3	30i	0V	0V	0V				

1. 60p、30i、30p和24p帧率也包含59.94p、29.97i和23.976p的帧率。
2.并不是EIAJ CP-4120规范的一部分，但设备通常都支持。
3. 相对于消隐电平。

98~99

5.8 其他专业视频模拟接口

表5-19和表5-20列出了其他的一些通用分量模拟视频格式。水平和垂直时序与525线（M）NTSC和625线（B、D、G、H、I）PAL是一样的。100%和75%彩条值见表5-21~表5-24。SMPTE、EBU N10、625线的Betacam和625线MII值与SDTV YPbPr是一样的。

表5-19 通用的专业视频分量模拟视频格式

格式	输出信号	信号幅度（V）	注释
SMPT、EBU N10	Y	+0.700	在Y上设置0%
	同步信号	−0.300	在饱和度上设置100%
	R′−Y，B′−Y	±0.350	三线为(Y+sync)、(R′−Y)、(B′−Y)
525线Betacam[1]	Y	+0.714	在Y上设置7.5%
	同步信号	−0.286	在饱和度上设置100%
	R′−Y，B′−Y	±0.467	三线为(Y+sync)、(R′−Y)、(B′−Y)
625线Betacam[1]	Y	+0.700	在Y上设置0%
	同步信号	−0.300	在饱和度上设置100%
	R′−Y，B′−Y	±0.350	三线为(Y+sync)、(R′−Y)、(B′−Y)
525线MII[2]	Y	+0.700	在Y上设置7.5%
	同步信号	−0.300	在饱和度上设置100%
	R′−Y，B′−Y	±0.324	三线为(Y+sync)、(R′−Y)、(B′−Y)
625线MII[2]	Y	+0.700	在Y上设置0%
	同步信号	−0.300	在饱和度上设置100%
	R′−Y，B′−Y	±0.350	三线为(Y+sync)、(R′−Y)、(B′−Y)

1. 索尼公司的商标。
2. 松下电器公司的商标。

表5-20　通用专业视频RGB模拟视频格式

格　式	输出信号	信号幅度（V）	注　释
SMPTE EBU N10	G′、B′、R′ 同步	+0.700 −0.300	在G′、B′和R′上设置0% 在饱和度上设置100% 三线为(G′+sync)、B′、R′
NTSC（设置）	G′、B′、R′ 同步	+0.714 −0.286	在G′、B′和R′上设置7.5% 在饱和度上设置100% 三线为(G′+sync)、B′、R′
NTSC（没有设置）	G′、B′、R′ 同步	+0.714 −0.286	在G′、B′和R′上设置0% 在饱和度上设置100% 三线为(G′+sync)、B′、R′
MII[1]	G′、B′、R′ 同步	+0.700 −0.300	在G′、B′和R′上设置7.5% 在饱和度上设置100% 三线为(G′+sync)、B′、R′

1. 松下电器公司的商标。

表5-21　525线Betacam 100%彩条，相对消隐电平的取值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	100	89.5	72.3	61.8	45.7	35.2	18.0	7.5
	mV	714	639	517	441	326	251	129	54
B′−Y	IRE	0	−65.3	22.0	−43.3	43.3	−22.0	65.3	0
	mV	0	−466	157	−309	309	−157	466	0
R′−Y	IRE	0	10.6	−65.3	−54.7	54.7	65.3	−10.6	0
	mV	0	76	−466	−391	391	466	−76	0

表5-22　525线Betacam 75%彩条，相对消隐电平的取值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	76.9	69.0	56.1	48.2	36.2	28.2	15.4	7.5
	mV	549	492	401	344	258	202	110	54
B′−Y	IRE	0	−49.0	16.5	−32.5	32.5	−16.5	49.0	0
	mV	0	−350	118	−232	232	−118	350	0
R′−Y	IRE	0	8.0	−49.0	−41.0	41.0	49.0	−8.0	0
	mV	0	57	−350	−293	293	350	−57	0

表5-23　525线MII 100%彩条，相对消隐电平的取值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	100	89.5	72.3	61.8	45.7	35.2	18.0	7.5
	mV	700	626	506	433	320	246	126	53
B′−Y	IRE	0	−46.3	15.6	−30.6	30.6	−15.7	46.3	0
	mV	0	−324	109	−214	214	−109	324	0
R′−Y	IRE	0	7.5	−46.3	−38.7	38.7	46.3	−7.5	0
	mV	0	53	−324	−271	271	324	−53	0

表5-24 525线MII 75%彩条，相对消隐电平的取值

		白色	黄色	青色	绿色	洋红色	红色	蓝色	黑色
Y	IRE	76.9	69.0	56.1	48.2	36.2	28.2	15.4	7.5
	mV	538	483	393	338	253	198	108	53
B′－Y	IRE	0	－34.7	11.7	－23.0	23.0	－11.7	34.7	0
	mV	0	－243	82	－161	161	－82	243	0
R′－Y	IRE	0	5.6	－34.7	－29.0	29.0	34.7	－5.6	0
	mV	0	39	－243	－203	203	243	－39	0

5.9 VGA接口

表5-25和图5-18给出了计算机设备和某些消费类设备所使用的、用于转换模拟RGB信号的15引脚VGA连接头。模拟RGB信号并不包含同步信息，也没有消隐基底电平，如图5-4所示。

表5-25 VGA连接头信号

引　脚	功　能	信号电平	阻　抗
1	红色	0.7V	75Ω
2	绿色	0.7V	75Ω
3	蓝色	0.7V	75Ω
4	接地		
5	接地		
6	接地－红色		
7	接地－绿色		
8	接地－蓝色		
9	＋5V DC		
10	接地－HSYNC（水平同步）		
11	接地－VSYNC（垂直同步）		
12	DDC SDA（数据）	≥2.4V	
13	HSYNC（水平同步）	≥2.4V	
14	VSYNC（垂直同步）	≥2.4V	
15	DDC SCL（时钟）	≥2.4V	

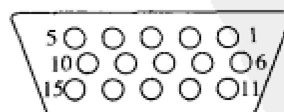


图5-18 VGA 15引脚D-SUB母连接头

参考文献

1. CEA-805 C, *Data on the Component Video Interfaces*, July 2006.
2. EIA-770.1 *Analog 525-Line Component Video Interface—Tree Channels*, November 2001.
3. EIA-770.2, *Standard-Definition TV Analog Component Video Interface*, November 2001.
4. EIA-770.3, *High-Definition IV Analog Component Video Interface*, November 2001.
5. EIA-J CPR-1204-1, *Transfer Method of Video ID Information Using Vertical Blanking Interval(525P System)*, March 1998.
6. EIA-J CPR-1204-2, *Transfer Method of Video ID Information Using Receiver Using D-Connector*, Jannuary 2000.

7. EIA-J CP-4120, *Interface Between Digital Tuner and Television Receiver Using D-Connector*, Jannuary 2000.
8. IEC 60933-1, Audio, *Video and Audiovisual Systems—Interconnections and Matching Values-Parts 1 : 21-pin Connector for Video Systems*, Application No.1, April 1988.
9. IEC 61880-2, *Video Systems (525/60)—Video and Accompanied Data Using the Vertical Blanking Interval—Part 2: 525 Progressive Scan System*, September 2002.
10. IEC 62375, *Video Systems (625/50 Progressive)—Video and Accompanied Data Using the Vertical Blanking Interval—Analog Interface*, February 2004.
11. ITU-R BT.709-5, 2002, *Parameter Values for the HDTV Standards for Production and International Programme Exchange.*
12. SMPTE 253M-1998, *Television-Three-Channel RGB Analog Video Interface.*
13. SMPTE 274M-2005, *Television-1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequence for Multiple Picture Rates.*
14. SMPTE 293M-2003, *Television—720 × 483 Active Line at 59.94Hz Progressive Scan Production-Digital Representation.*
15. SMPTE RP-160-1997, *Tree-Channel Parallel Analog Component High-Definition Video Interface.*

100 ~ 105



第6章　数字视频接口

6.1　专业视频分量接口

专业视频设备具有一些特定要求，比如在演播室内部使用的这些专业视频设备，因此，它们具有自身的一组数字视频互连标准。表6-1列出了用于各种专业视频的并行和串行数字接口标准。

表6-1　专业视频各种分量视频格式的并行和串行数字接口标准。i=隔行，p=逐行

有效分辨率（H×V）	总的分辨率（H×V）	显示器宽高比	帧率（Hz）	1×Y采样率（MHz）	SDTV还是HDTV	数字并行标准	数字串行标准
720×480i	858×525i	4：3	29.97	13.5	SDTV	BT.656 BT.799 SMPTE 125M	BT.656 BT.799
720×480p	858×525p	4：3	59.94	27	SDTV	–	BT.1362 SMPTE 294M
720×576i	864×625i	4：3	25	13.5	SDTV	BT.656 BT.799	BT.656 BT.799
720×576p	864×625p	4：3	50	27	SDTV	–	BT.1362
960×480i	1144×525i	16：9	29.97	18	SDTV	BT.1302 BT.1303 SMPTE 267M	BT.1302 BT.1303
960×576i	1152×625i	16：9	25	18	SDTV	BT.1302 BT.1303	BT.1302 BT.1303
1280×720p	1650×750p	16：9	59.94	74.176	HDTV	SMPTE 274M	–
1280×720p	1650×750p	16：9	60	74.25	HDTV	SMPTE 274M	–
1920×1080i	2200×1125i	16：9	29.97	74.176	HDTV	BT.1120SMPTE 274M	BT.1120 SMPTE 292M
1920×1080i	2200×1125i	16：9	30	74.25	HDTV	BT.1120 SMPTE 274M	BT.1120 SMPTE 292M
1920×1080p	2200×1125p	16：9	59.94	148.35	HDTV	BT.1120SMPTE 274M	–
1920×1080p	2200×1125p	16：9	60	148.5	HDTV	BT.1120SMPTE 274M	–
1920×1080i	2376×1250i	16：9	25	74.25	HDTV	BT.1120	BT.1120
1920×1080p	2376×1250p	16：9	50	148.5	HDTV	BT.1120	–

6.1.1　视频时序

处理消隐时，并不是将消隐期数字化然后进行传输，而是在数字视频流内插入一些特定的序列来标明有效视频的起始（SAV）和结束（EAV）。这些SAV和EAV序列指明什么时候提供水平和场消隐，还指明了需要传输哪个场（field）。这样就可以在消隐期内传输像数字音频、图文电视、字幕等这些辅助数据。

EAV和SAV序列具有的优先级必须高于有效视频数据或辅助数据，以保证接收机总可以维持正确的视频时序。

编码器的视频时序序列是由第4章已经讨论过的三个时序信号来控制的，这三个时序信号分别为：H（行消隐）、V（场消隐）和F（场1或场2）。H由0跃迁到1就会触发EAV序列，而由1跃迁到0就会触发SAV序列。F和V只能在EAV序列处发生改变。

通常，专业视频分量接口可以同时支持8位和10接口，对于10位接口，其中2位用于传输视频数据的小数部分，这样可以尽可能地降低累积处理误差，并且可以支持10位的辅助数据。

106～107

YCbCr和R′G′B′数据可以不使用10位数中位于0x000~0x003和0x3FC~0x3FF范围内的值。或不使用8位数中的0x00或0xFF这两个值，因为需要将它们用作时序信息。

EAV和SAV序列见表6-2。状态字定义为：

F＝“0”表示场1　　　　F＝“1”表示场2

V＝“1”（在场消隐期）

H＝“0”（处于SAV时）　　　　H＝“1”（处于EAV时）

P3~P0＝保护位

P3＝V⊕H

P2＝F⊕H

P1＝F⊕V

P0＝F⊕V⊕H

其中⊕表示异或操作。有了这些保护位，在接收端就能够检测出1~2位的错误码，并且能够检测并纠正1位错误码。对于大多数逐行显示的视频系统，由于没有场信息，F通常为“0”。

表6-2　EAV和SAV序列

	8位数据								10位数据	
	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
前导	1	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0
状态字	1	F	V	H	P3	P2	P1	P0	0	0

对于4：2：2的YCbCr数据，在每个SAV序列之后，有效数据字（word）流总是从Cb采样点开始，如图6-1所示。在复用序列中，协调位置采样点（对应于图像上同一点的那些采样点）被分为组Cb、Y和Cr。在消隐期内，如果没有提供辅助数据，那么10位的Y值或R′G′B′值应该设为0x40，而10位的CbCr值应该设为0x200。

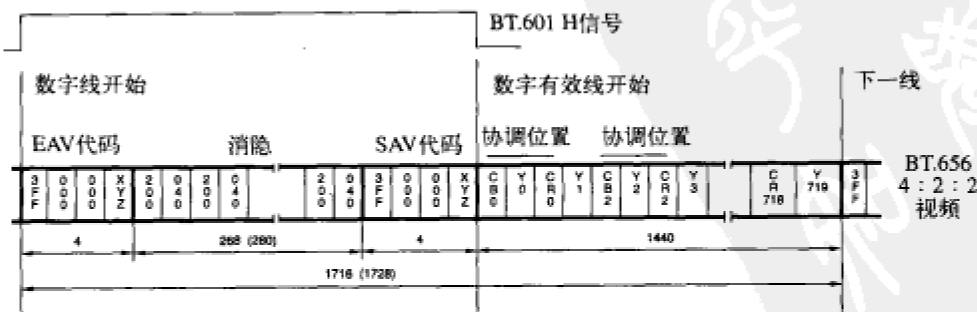


图6-1　BT.656并行接口的一行扫描线数据。480i，4：2：2 YCbCr，每线有720个有效采样点，27MHz时钟，10位系统。576i系统的值在括号内给出

接收设备通过查找8位的0xFF 0x00 0x00前导来检测EAV和SAV序列。状态字（在接收设备端进行误码改正，这是可选的，见表6-3）用于恢复H、V和F时序信号。

表6-3 在解码器中进行SAV和EAV误码纠正

接收到的F、V、H(位D8~D6) 接收的D5~D2	000	001	010	011	100	101	110	111
0000	000	000	000	*	000	*	*	111
0001	000	*	*	111	*	111	111	111
0010	000	*	*	011	*	101	*	*
0011	*	*	010	*	100	*	*	111
0100	000	*	*	011	*	*	110	*
0101	*	001	*	*	100	*	*	111
0110	*	011	011	011	100	*	*	011
0111	100	*	*	011	100	100	100	*
1000	000	*	*	*	*	101	110	*
1001	*	001	010	*	*	*	*	111
1010	*	101	010	*	101	101	*	101
1011	010	*	010	010	*	101	010	*
1100	*	001	110	*	110	*	110	110
1101	001	001	*	001	*	001	110	*
1110	*	*	*	011	*	101	110	*
1111	*	001	010	*	100	*	*	*

注：* = 不可纠正的错误。

6.1.2 辅助数据

辅助数据包用于在消隐期内发送非视频数据信息（如数字音频、隐藏字幕和图文电视等）。广泛使用的各种ITU-R和SMPTE规范描述了各种辅助数据格式。

在水平消隐期内，辅助数据可以在EAV和SAV序列之间的间隔内传输。在场消隐期内，辅助数据可以在SAV和EAV序列之间的间隔内传输。在水平和场消隐期内可以提供多个辅助数据包，但这些辅助数据包之间必须是连续的。

108
109

有两种类型的辅助数据格式。较早的Type 1格式使用单个数据ID字标明辅助数据的类型，而较新的Type 2格式使用2个字来表示数据ID。表6-4给出通用的包格式。

表6-4 辅助数据包通用格式

	8位数据								10位数据	
	D9（最高有效位）	D8	D7	D6	D5	D4	D3	D2	D1	D0
辅助数据标识（ADF）	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1
数据ID（DID）	$\overline{D8}$	偶校验位	从0000 0000到1111 1111的值							
数据块编号或SDID	$\overline{D8}$	偶校验位	从0000 0000到1111 1111的值							
数据计数（DC）	$\overline{D8}$	偶校验位	从0000 0000到1111 1111的值							
用户数据字0	从00 0000 0100 到 11 1111 1011的值									
	⋮									
用户数据字N	值00 0000 0100 到 11 1111 1011									
校验和	$\overline{D8}$		通过最后的用户数据字，D0~D8表示的数据ID的和 预设为全0，内容被忽略							

1. 数据ID（DID）

DID表示所发送数据的类型。大多数DID值的指定是由ITU和SMPTE控制的，这是为了保证设备的兼容性。新DID值在不需要注册的情况下也是可用的。

2. 辅助ID（SDID，仅适用于Type 2）

对于Type 2附加格式，SDID是数据ID的一部分。为了保证设备的兼容性，大多数SDID值的指定也是由ITU和SMPTE控制的。新的SDID值在不需要注册的情况下也是可用的。

3. 数据块编号（DBN，仅适用于Type 1）

使用DBN，在接收端就可以将多个附加包（共用相同的DID）通过延迟等待的方式收集在一起。当需要发送的用户数据字数目多于255时，就会遇到这种情形，因此需要一个以上的辅助数据包来发送这些数据。对于每个连续的辅助数据包，DBN值增加1。

4. 数据计数（DC）

DC指定了包中用户数据字的数目。在8位的应用系统中，DC指定了6个8位值的MSB（最高有效位），因此，用户数据字的数目必须为4的整倍数。

5. 用户数据字（UDW）

在包中可以提供最多255个用户数据字。在8位的应用系统中，用户数据字必须为4的整倍数。可能需要填充一些额外的数据字，以确保提供的用户数据为4的整倍数。

用户数据可以不使用10位值0x000~0x003和0x3FC~0x3FF，或8位值0x00和0xFF，因为它们要用作时序信息。

6.1.3　并行接口

1. 25引脚并行接口

这类接口主要是用于传输SDTV分辨率、格式为4：2：2的YCbCr数据。传输8位或10位数据和时钟。单独的数据位用D0~D9表示，其中D9为最高有效位。信号的引脚分配见表6-5。

表6-5　25引脚并行接口连接头引脚的分配。8位接口使用 D9~D2

引　脚	信　号	引　脚	信　号
1	时钟	14	时钟-
2	系统接地A	15	系统接地B
3	D9	16	D9-
4	D8	17	D8-
5	D7	18	D7-
6	D6	19	D6-
7	D5	20	D5-
8	D4	21	D4-
9	D3	22	D3-
10	D2	23	D2-
11	D1	24	D1-
12	D0	25	D0-
13	线缆屏蔽		

Y有10个数据位，标称表示范围为0x040~0x3AC。小于0x040或大于0x3AC的值由于处理的原因也需要提供。在消隐期内，如果没有提供数据，那么数据Y的值应该为040_H。

Cb和Cr有10个数据位，标称表示范围为0x040~0x3C0。由于处理的原因，小于0x040和大于0x3C0的值也可能需要提供。在消隐期内，如果没有其他的数据可提供，那么CbCr数据的值应该

为0x200。

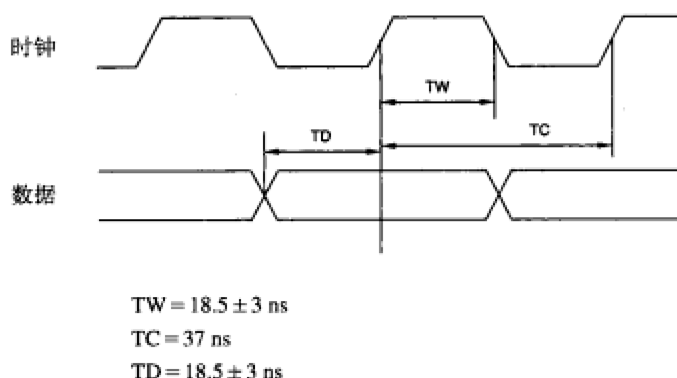
对于ECL兼容的平衡驱动设备和接收设备（balanced driver and receiver），信号电平是兼容的。信号生成器必须具有最大源阻抗为110Ω的平衡输出；信号通过110Ω负载时测得的峰间电压必须在0.8~2.0V之间。在接收设备端，传输线路终止于110±10Ω的负载。

110~111

❑ 27MHz并行接口

这种BT.656和SMPTE 125M接口用于宽高比为4：3的480i和576i系统。采样率为13.5MHz的Y信息和复用CbCr信息，采用复用方式，以27MHz时钟频率形成一个单一的8位或10位数据流。

27MHz时钟信号的时钟脉冲宽度为18.5±3ns。时钟信号的正跃变发生在限差为±3ns的数据跃变中点（如图6-2所示）。



TW = 18.5 ± 3 ns
TC = 37 ns
TD = 18.5 ± 3 ns

图6-2 25引脚27MHz并行接口波形

为了可以在连接长度为50~200米的范围内进行可靠的操作，接收端必须使用如图6-3所示的典型特征进行频率均衡。在本例中，线缆长度在直到0的各种长度范围内都可以实现可靠操作。

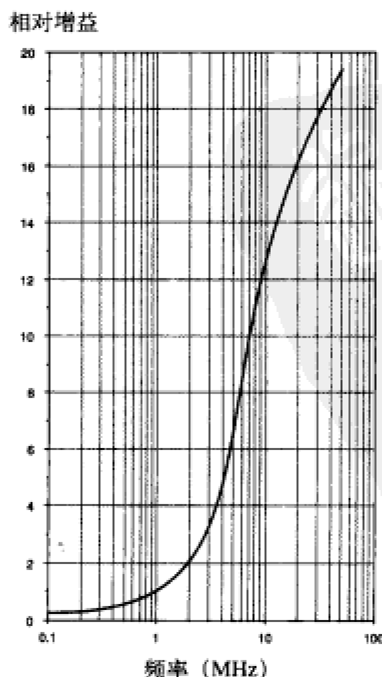
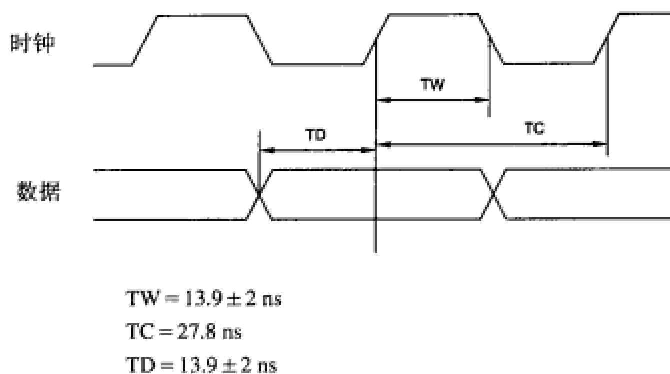


图6-3 针对小信号的线性接收设备均衡特征示例

❑ 36MHz并行接口

这种BT.1302和SMPTE 267M接口被用于宽高比为16：9的480i和576i系统。采样频率为18MHz的Y信息和复用CbCr信息，以36MHz的时钟频率，采用复用的方式形成一个单一的8位或10位数据流。

时钟频率为36MHz的时钟脉冲宽度为13.9±2ns。时钟信号的正跃变发生在限差为±2ns的数据跃变中点（如图6-4所示）。



TW = 13.9 ± 2 ns
TC = 27.8 ns
TD = 13.9 ± 2 ns

图6-4　25引脚36MHz并行接口波形

112~113

为了在40~160m连接长度范围内可以进行可靠的操作，接收端必须使用典型特征进行频率均衡，见图6-3。在本例中，线缆长度在直到0的各种长度范围内都可以实现可靠操作。

2. 93引脚并行接口

这种接口是用于对HDTV分辨率的R′G′B′数据、4：2：2的YCbCr数据或4：2：2：4的YCbCrK数据进行变换。这些信号的引脚分配如表6-6所示。最高有效位为R9、G9和B9。

表6-6　93引脚并行接口的连接头引脚分配。8位接口只使用了9~2位

引脚	信号	引脚	信号	引脚	信号	引脚	信号
1	时钟	20	GND	39	G4-	58	R5
2	G9	21	GND	40	G3-	59	R4
3	G8	22	GND	41	G2-	60	R3
4	G7	23	GND	42	G1-	61	R2
5	G6	24	GND	43	G0-	62	R1
6	G5	25	GND	44	B9-	63	R0
7	G4	26	GND	45	B8-	64	GND
8	G3	27	GND	46	B7-	65	GND
9	G2	28	GND	47	B6-	66	GND
10	G1	29	GND	48	B5-	67	GND
11	G0	30	GND	49	B4	68	GND
12	B9	31	GND	50	B3	69	GND
13	B8	32	GND	51	B2	70	GND
14	B7	33	时钟-	52	B1	71	GND
15	B6	34	G9-	53	B0	72	GND
16	B5	35	G8-	54	R9	73	GND
17	GND	36	G7-	55	R8	74	GND
18	GND	37	G6-	56	R7	75	GND
19	GND	38	G5-	57	R6	76	GND

（续）

引脚	信号	引脚	信号	引脚	信号	引脚	信号
77	GND	83	B0-	89	R4-		
78	GND	84	R9-	90	R3-		
79	B4-	85	R8-	91	R2-		
80	B3-	86	R7-	92	R1-		
81	B2-	87	R6-	93	R0-		
82	B1-	88	R5-				

当传输4：2：2：4 YCbCr数据时，绿色通道加载的是Y信息，红色通道加载的是复用的CbCr信息。

当传输4：2：2：4 YCbCrK数据时，绿色通道加载的是Y信息，红色通道加载的是复用的CbCr信息，绿色通道加载的是K（alpha通道）信息。

Y是一个10位的数据，标称表示范围为0x040~0x3AC。由于处理的原因，也可能提供小于040_H或大于0x3AC的值。在消隐期内，如果没有提供其他的数据，数据Y的值应该为0x040。

Cb和Cr是10位的数据，标称表示范围为0x040~0x3C0。由于处理的原因，也可以提供小于0x040或大于0x3C0的值。在消隐期内，如果没有提供其他的信息，那么CbCr数据的值应该为0x200。

R′G′B′和K是10位的数据，标称表示范围为0x040~0x3AC。由于处理的原因，也可以提供小于0x040和大于0x3AC的值。在消隐期内，如果没有提供其他的信息，那么R′G′B′数据的值应该为0x040。

对于ECL兼容的平衡驱动设备和接收设备，信号电平是兼容的。信号发生器必须是平衡输出，输出最大源阻抗为110Ω；信号必须为0.6~2.0V，这是跨110Ω负载测得的峰间电压值。在接受端，传输线路必须终止于110±10Ω的负载。

❑ 74.25和74.176 MHz并行接口

这种ITU-R BT.1120和SMPTE 274M接口主要用于HDTV系统。

74.25或74.176MHz（74.25/1.001）时钟信号的时钟脉冲宽度为6.73±1.48ns。时钟信号的正跃迁发生在数据跃迁之间的中点，数据跃迁的限差（误差容限）为±1ns（如图6-5所示）。

114~115

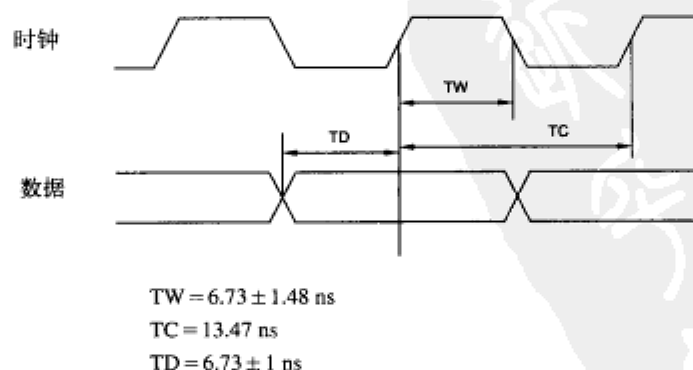


图6-5 93引脚74.25和74.176 MHz并行接口波形

为了在互连长度大于20m时能够进行可靠的操作，接收端必须使用频率均衡。

❑ 148.5和148.35 MHz并行接口

这种BT.1120和SMPTE 274M接口用于HDTV系统。

148.5或148.35MHz（148.5/1.001）时钟信号的脉冲宽度为3.37±0.74ns。时钟信号的正跃迁发生在数据跃迁的中点，数据跃迁的限差为±0.5ns（类似于图6-5）。

为了在大于14m的互连长度上能够进行可靠的操作，接收端必须进行频率均衡。

3. 应用

可以用一个或多个并行接口在设备之间传输各种视频格式。

❑ 4：2：2 YCbCr-隔行SDTV

ITU-R BT.656和BT.1302并行接口是用于在设备之间传输BT.601 4：2：2 YCbCr数字视频的。SMPTE 125M和267M进一步明确了针对480i系统的操作。

图6-6表示宽高比为4：3、采样时钟频率为27MHz时，一行扫描线的时序。图6-7表示的是宽高比为16：9、采样时钟频率为36MHz时，一行扫描线的时序。4：2：2YCbCr-隔行SDTV使用的是25引脚并行接口。

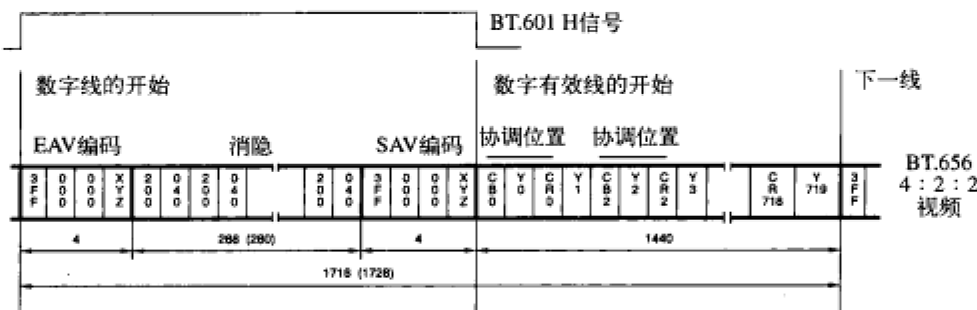


图6-6　针对一行扫描线的BT.656和SMPTE 125M并行接口数据。480i；4：2：2 YCbCr；每线有720有效采样点；27MHz时钟频率；10位系统。576i系统的值在括号内给出

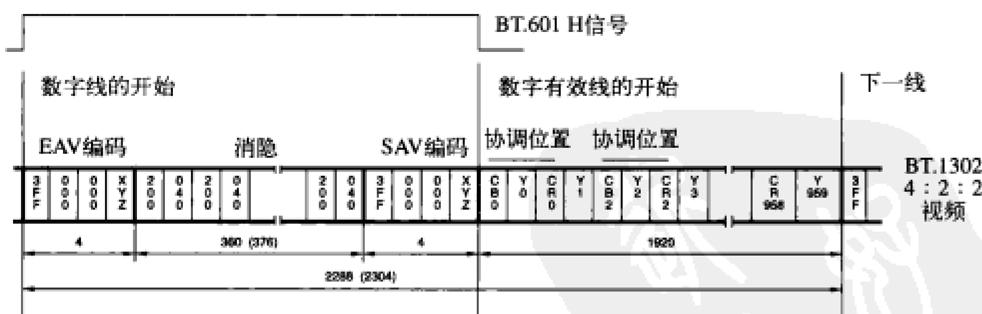


图6-7　针对一行扫描线的BT.1302和SMPTE 267M并行接口数据。480i；4：2：2 YCbCr；每线有960有效采样点；36MHz时钟频率；10位系统。576i系统的值在括号内给出

❑ 4：4：4 YCbCr-隔行SDTV

ITU BT.799和BT.1303并行接口用于在设备之间传输BT.601的4：4：4的YCbCrK数字视频。K为alpha信号，用于混合不同的视频源，将在第7章对此进行讨论。SMPTE RP-175更进一步明确了480i系统的操作。

116
117

使用了两种传输链。链A包含所有的Y采样点和那些位于偶数采样点的Cb和Cr采样点。链B包含抠像通道（keying channel）和那些位于奇数采样点位置的Cb和Cr采样点。尽管通常可以称链A为4：2：2，链B为2：2：4，但是链A并不是真正的4：2：2信号，因为CbCr数据的采样频率为13.5MHz，而不是6.75MHz。

图6-8表示的是在传输4：4：4：4YCbCrK视频数据时，链A和链B的内容。图6-9表示的是在传输R′G′B′K视频数据时，链A和链B的内容。如果没有提供alpha值（K），那么K采样点值应该是10位的$3AC_H$值。

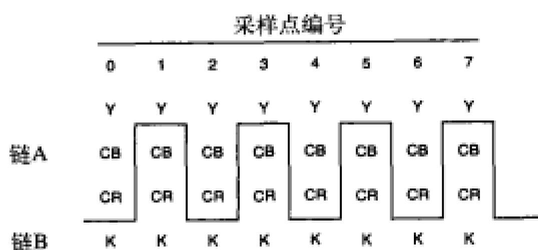


图6-8 YCbCrK视频信号的链内容表示



图6-9 R′G′B′K视频数据的链内容表示

图6-10表示的是宽高比为4：3、采样时钟频率为27MHz时，一行扫描线的时序。图6-11表示的是宽高比为16：9、采样时钟频率为36MHz时，一行扫描线的时序。4：4：4YCbCr-隔行SDTV使用的是两个25引脚的并行接口。

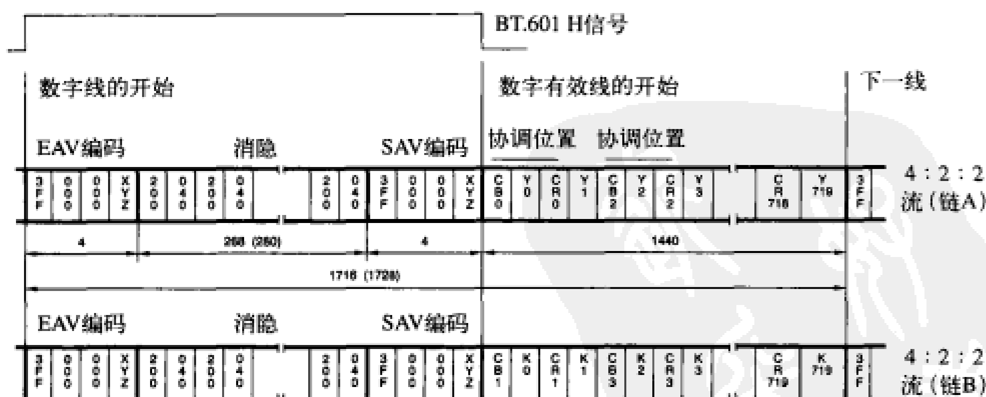


图6-10 针对一行扫描线的BT.799和SMPTE RP-175并行接口数据。480i；4：4：4：4 YCbCrK；每线有720有效采样点；27MHz时钟频率；10位系统。576i系统的值在括号内给出

❑ RGBK-隔行SDTV

BT.799和BT.1303也支持在设备间传输BT.601R′G′B′K数字视频数据。如果要了解更多地相关信息，参见4：4：4：4 YCbCr接口。SMPTE RP-175进一步明确了480i系统的操作。G′采样点数据是在Y位置发送的，R′采样点数据是在Cr位置发送的，而B′采样点数据是在Cb位置发送的。

❑ 4：2：2 YCbCr- 逐行SDTV

ITU-R BT.1362定义了两个10位的4：2：2 YCbCr数据流（见图6-12），使用的时钟频率为

27MHz。SMPTE 294M更进一步地明确了480p系统的操作。在表6-7中给出了各个扫描线使用的流。

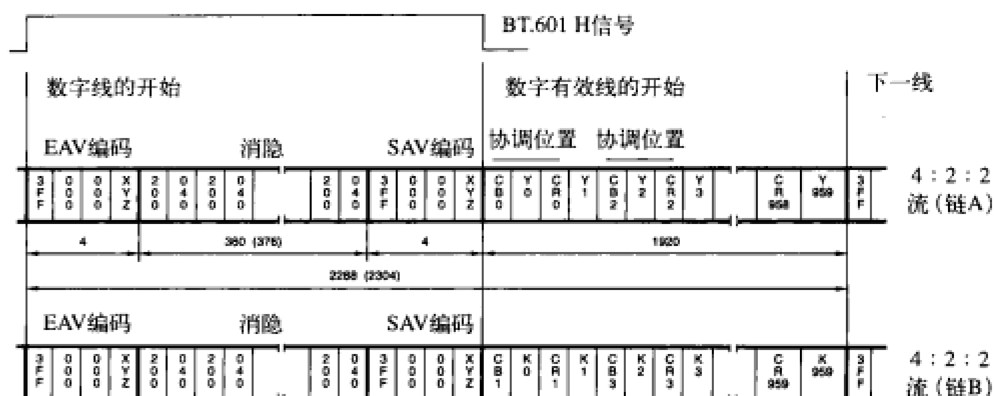


图6-11　针对一行扫描线的BT.1303并行接口数据。480i，4：4：4：4 YCbCrK；每线有960有效采样点；36MHz时钟频率；10位系统。576i系统的值在括号内给出

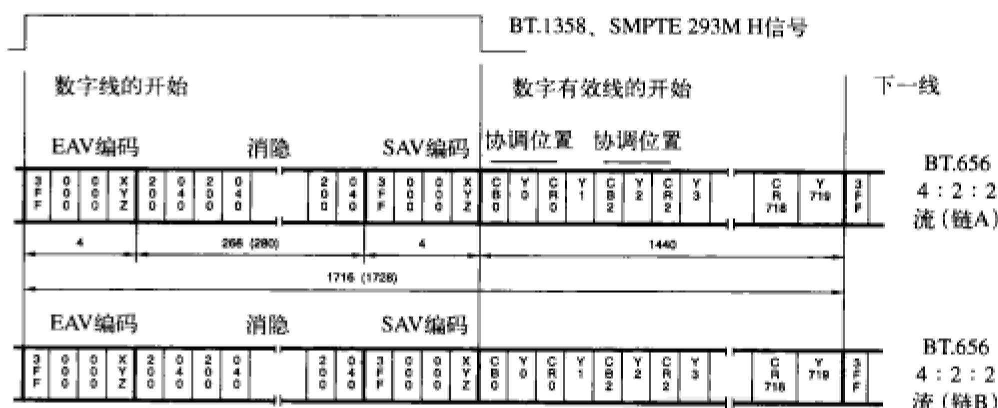


图6-12　针对一行扫描线的BT.1362和SMPTE 294M并行接口数据。480p；4：2：2 YCbCr；每线有720有效采样点；27 MHz时钟频率；10位系统。576i系统的值在括号内给出

表6-7　BT.1362和SMPTE 294M扫描线编号和链分配

480p（525p）系统				576p（625p）系统			
链A	链B	链A	链B	链A	链B	链A	链B
7	8	6	7	1	2	4	5
9	10	⋮	⋮	3	4	6	7
⋮	⋮	522	523	⋮	⋮	8	9
523	524	524	525	621	622	⋮	⋮
525	1	1	2	623	624	620	621
2	3	3	4	625	1	622	623
4	5	5	6	2	3	624	625

❑ 4：2：2 YCbCr-隔行HDTV

ITU-R BT.1120并行接口是用于在设备间传输隔行HDTV 4：2：2 YCbCr数字视频数据的。SMPTE 274M更进一步明确了29.97Hz和30Hz系统的操作。

图6-13表示的是针对1920×1080i有效分辨率一行扫描线的时序。93引脚并行接口可以采用74.25MHz（25Hz或30Hz的帧率）或74.176MHz（29.97Hz的帧率）的采样时钟频率。

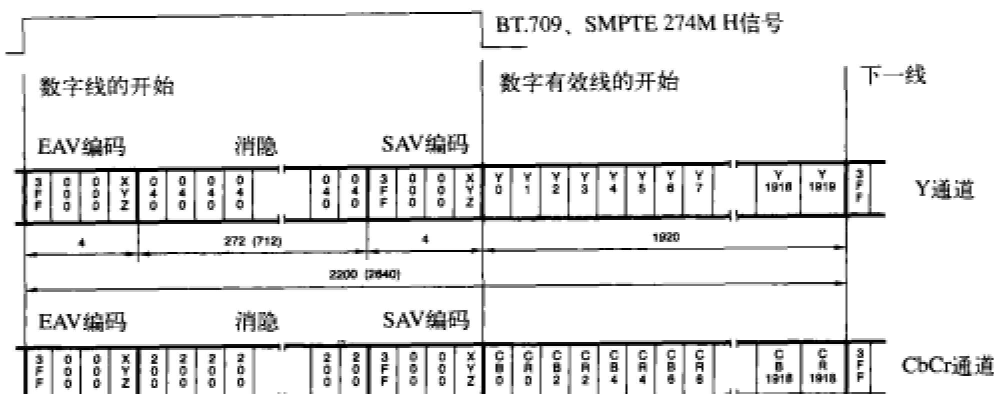


图6-13 针对一行扫描线的BT.1120和SMPTE 274M并行接口数据。1080i29.97、1080i30、1080p59.94和1080p60系统；4：2：2 YCbCr；每线有1920有效采样点；时钟频率为74.176、74.25、148.35或148.5 MHz；10位系统。1080i25和1080p50系统的值在括号内给出

❑ 4：2：2：4 YCbCrK-隔行扫描HDTV

BT.1120也支持在设备间传输HDTV的4：2：2：4 YCbCrK数字视频数据。SMPTE 274M进一步明确了29.97Hz和30Hz系统的操作。

图6-14表示的是针对1920×1080i有效分辨率一行扫描线的时序。93引脚并行接口可用的采样时钟频率为74.25MHz（25或30Hz帧率）或74.176MHz（29.97Hz帧率）。

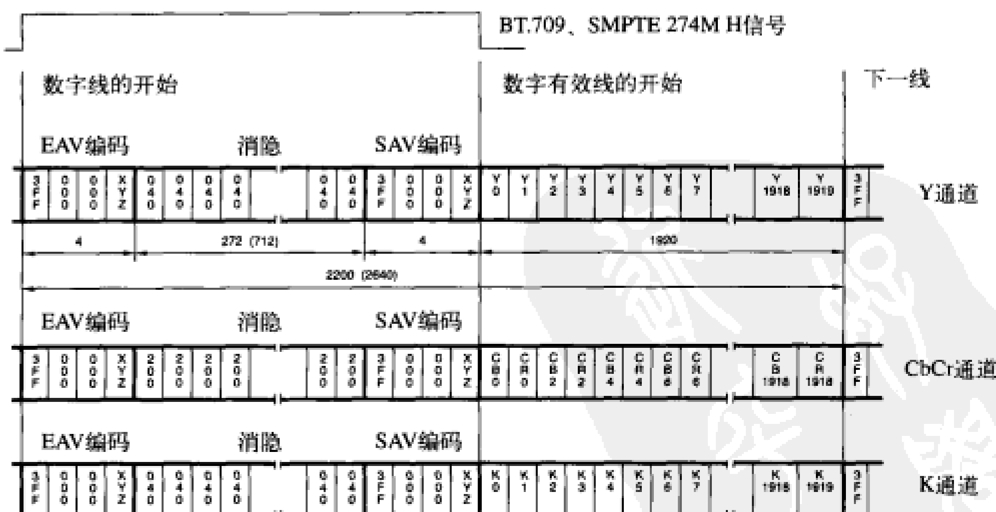


图6-14 针对一行扫描线的BT.1120和SMPTE 274M并行接口数据。1080i29.97、1080i30、1080p59.94和1080p60系统；4：2：2：4 YCbCrK；每线有1920有效采样点；时钟频率为74.176、74.25、148.35或148.5MHz；10位系统。1080i25和1080p50系统的值在括号内给出

❑ RGB-隔行HDTV

BT.1120也支持在设备间传输HDTV的R′G′B′数字视频。SMPTE 274M更进一步明确了29.97Hz和30Hz系统的操作。

图6-15表示的是针对1920×1080i有效分辨率一行扫描线的时序。93引脚并行接口可用的采样

118 ~ 123

时钟频率为74.25MHz（25或30Hz帧率）或74.176MHz（29.97Hz帧率）。

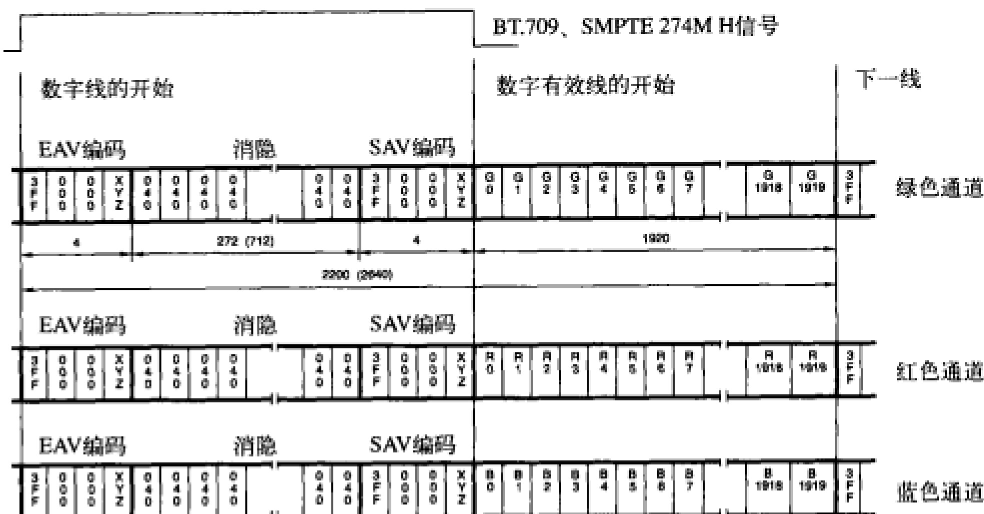


图6-15　针对一行扫描线的BT.1120和SMPTE 274M并行接口数据。1080i29.97、1080i30、1080p59.94和1080p60系统；R'G'B'；每线有1920有效采样点；时钟频率为74.176、74.25、148.35或148.5 MHz；10位系统。1080i25和1080p50系统的值在括号内给出

❑ 4：2：2 YCbCr-逐行HDTV

ITU-R BT.1120和SMPTE 274M并行接口是用于在设备间传输逐行HDTV 4：2：2 YCbCr数字视频的。

图6-13表示的是针对1920×1080i有效分辨率一行扫描线的时序。93引脚并行接口使用的采样时钟频率为148.5MHz（24Hz、25Hz、30Hz、50Hz或60Hz的帧率）或148.35MHz（23.98Hz、29.97Hz或59.94Hz的帧率）。

图6-16表示的是针对1280×720p有效分辨率一行扫描线的时序。93引脚并行接口使用的采样时钟频率为74.25MHz（24Hz、25Hz、30Hz、50Hz或60Hz的帧率）或74.176MHz（23.98Hz、29.97Hz或59.94Hz的帧率）。

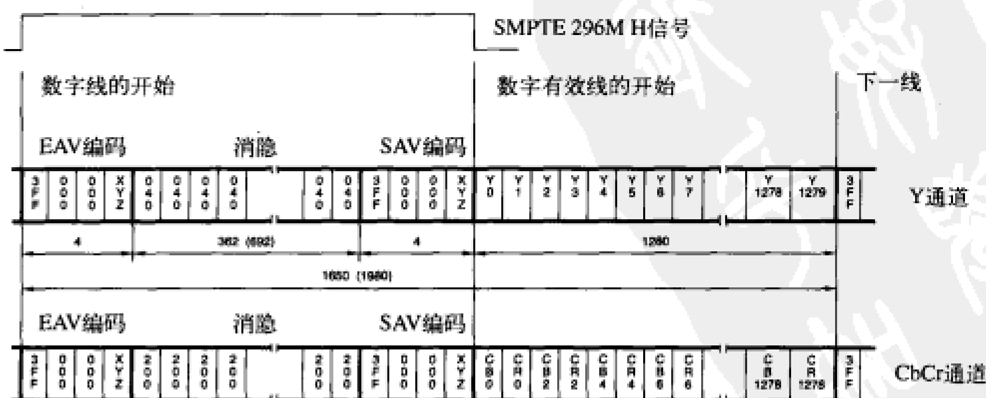


图6-16　SMPTE 274M针对一行扫描线的并行接口数据。720p 59.94和720p60系统；4：2：2 YCbCr；每线1280个有效采样点；74.176或74.25 MHz时钟频率；10位系统。720p50系统的值在括号内给出

❑ 4：2：2：4 YCbCrK-逐行HDTV

BT.1120和SMPTE 274M也支持在设备间传输HDTV 4：2：2：4 YCbCr数字视频。

图6-14表示的是1920×1080p有效分辨率一行扫描线的时序。93引脚并行接口使用的时钟频率为148.5MHz（24、25、30、50和60Hz的帧率）或148.35MHz（23.98、29.97或59.94Hz帧率）。

图6-17表示的是针对1280×720p有效分辨率一行扫描线的时序。93引脚并行接口使用的时钟频率为74.25MHz（24、25、30、50和60Hz的帧率）或74.176MHz（23.98、29.97或59.94Hz帧率）。

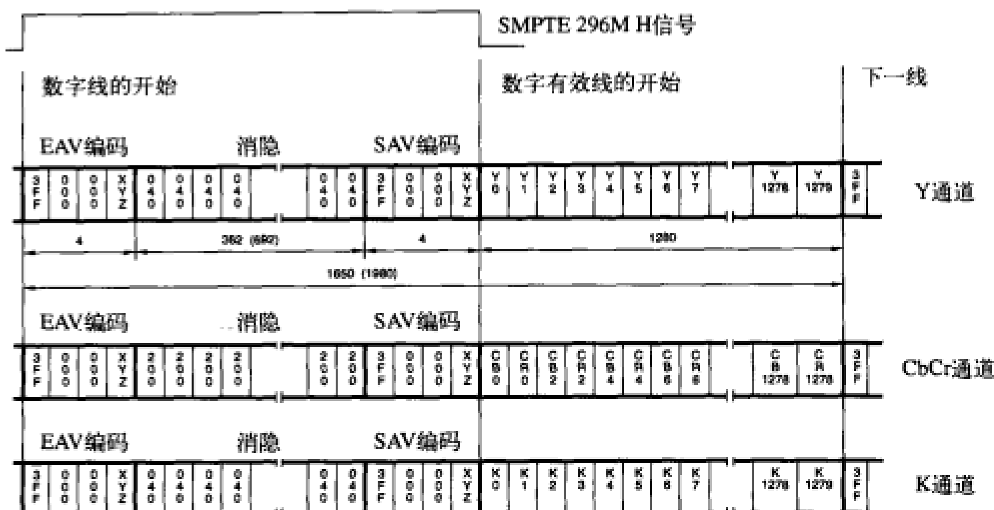


图6-17 SMPTE 274M针对一行扫描线的并行接口数据。720p 59.94和720p60系统，4：2：2：4 YCbCrK，每线1280个有效采样点，74.176或74.25MHz时钟频率，10位系统。720p50系统的值在括号内给出

❑ RGB-逐行HDTV

BT.1120和SMPTE 274M也支持在设备间传输HDTV R′G′B′数字视频。

图6-15表示的是针对1920×1080p有效分辨率一行扫描线的时序。93引脚并行接口使用的时钟频率为148.5MHz（24、25、30、50和60Hz的帧率）或148.35MHz（23.98、29.97或59.94Hz帧率）。

图6-18表示的是针对1280×720p有效分辨率一行扫描线的时序。93引脚并行接口使用的时钟频率为74.25MHz（24、25、30、50和60Hz的帧率）或74.176MHz（23.98、29.97或59.94Hz帧率）。

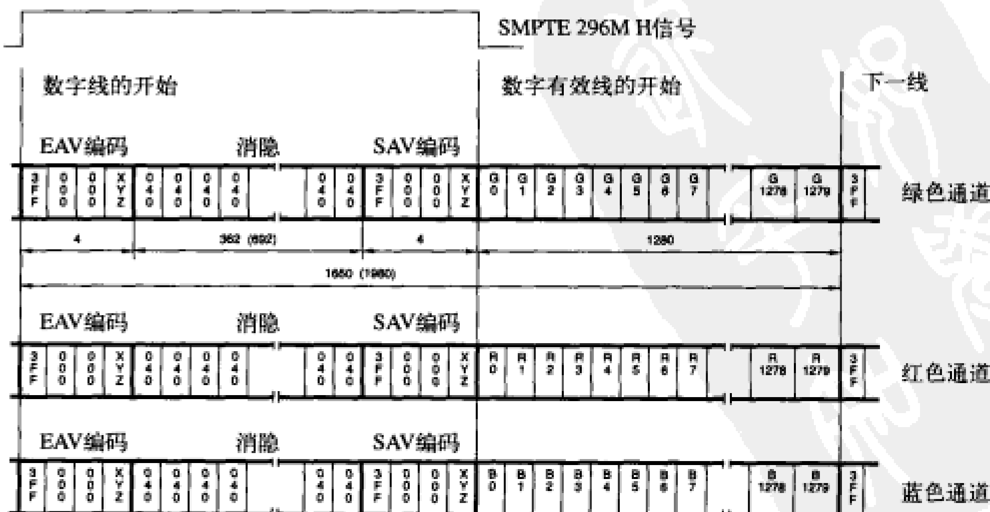


图6-18 SMPTE 274M针对一行扫描线的并行接口数据。720p 59.94和720p60系统，R′G′B′，每线1280个有效采样点，74.176或74.25 MHz时钟频率，10位系统。720p50系统的值在括号内给出

6.1.4　串行接口

并行格式可以转换为串行格式（见图6-19），使得数据可以在75Ω的同轴有线或光纤中传输。

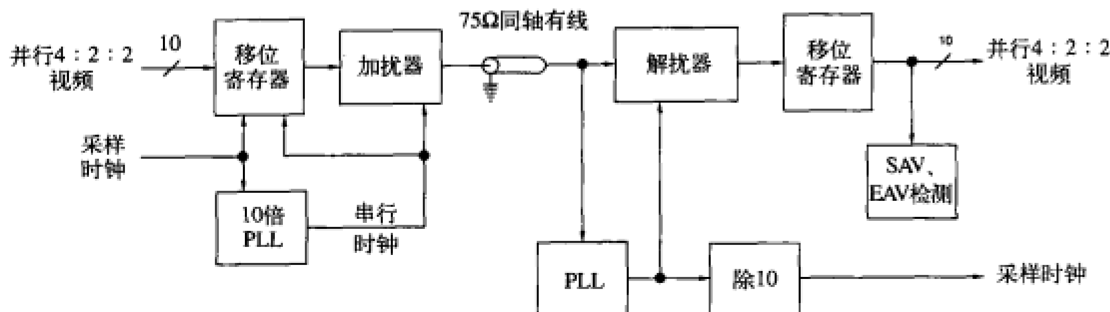


图6-19　串行接口框图

对于有线连接方式，信号产生器具有不平衡的输出，其源阻抗为75Ω；信号在75Ω负载间测得的峰间电压必须为0.8V±10%。接收端的输出阻抗为75Ω。

在8位系统的环境中，串行化之前，在EAV和SAV期间，编码0x00和0xFF分别被扩展为10位的值0x000和0x3FF。而所有其他的8位数据在串行化之前都要在最低有效位后附加两个0。

对10位数据进行串行化（LSB优先），并使用加扰和不分极性NRZI算法进行处理：

$G(x)=(x^9+x^4+1)(x+1)$

输入到加扰器的信号（见图6-20）采用正逻辑（最高电压表示逻辑1，最低电压表示逻辑0）。

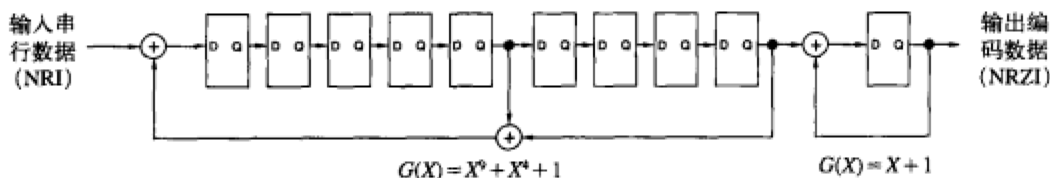


图6-20　典型的加扰电路

124～127

格式化的串行数据以10倍的采样时钟频率输出。由于并行时钟可能包含大量抖动（jitter），因此直接由没有经过滤波的并行时钟推导得到的10倍采样时钟频率可能导致过大的信号扰动。

在接收端，锁相同步（phase-lock synchronization）是通过检测EAV和SAV序列来完成的。PLL不断地微调每行扫描线来确保能够检测到这些模式，并且避免发生跳位（bit slippage）。尽管必须慎重，以避免掩盖与字相关的扰动成分，但恢复的10倍采样时钟还需要除以10来生成采样时钟频率。串行数据还要进行高低频率均衡、解扰（见图6-21）和反序列化。

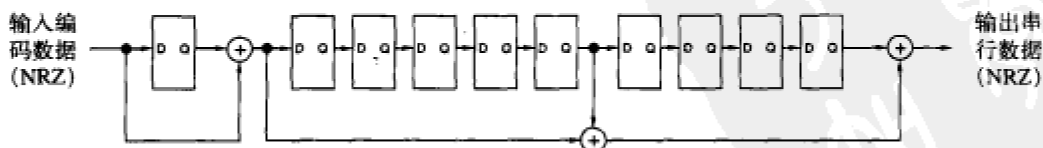


图6-21　典型的解扰器电路

1. 270 Mbit/s串行接口

这种BT.656和SMPTE 259M接口（也称为SDI）是用于将27MHz的并行数据流转换为270Mbit/s的串行数据流。10倍PLL使用27MHz时钟信号来生成270MHz的时钟。这种接口主要用于

480i和576i的4：3系统。

2. 360 Mbit/s串行接口

这种BT.1302和SMPTE 259M接口是用于将36MHz并行数据流转换为360Mbit/s的串行数据流。10倍PLL使用36MHz时钟信号生成360MHz的时钟。这种接口主要用于480i和576i的16：9系统。

3. 540 Mbit/s串行接口

这种SMPTE 344M接口是用于将54MHz或27MHz并行数据流转换为540Mbit/s的串行数据流。10倍PLL使用54MHz时钟信号生成540MHz的时钟。这种接口主要用于宽高比为4：3的480p和576p系统。

4. 1.485和1.4835 Gbit/s串行接口

这种BT.1120和SMPTE 292M接口是用于将两路74.25MHz或74.176（74.25/1.001）MHz并行数据流（Y和CbCr），以复用的方式生成单一的1.485Gbit/s或1.4835Gbit/s的串行数据流。20倍PLL使用74.25或74.176MHz时钟信号生成1.485GHz的时钟。这种接口主要用于HDTV系统。

在以复用的方式合并两路并行数据流之前，需要将行数和CRC信息（见表6-8）加入每个流的每个EAV流之后。CRC用于检测有效视频和EAV中的错误。它是由如下多项式生成的两个字组成：

128

$CRC = x^{18} + x^5 + x^4 + 1$

其初始值设为0。计算过程开始于第一个有效线字（active line word），结束于线编号（LN1）的最后一个字。

表6-8 线编号和CRC数据

	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
LN0	$\overline{D8}$	L6	L5	L4	L3	L2	L1	L0	0	0
LN1	$\overline{D8}$	0	0	0	L10	L9	L8	L7	0	0
CRC0	$\overline{D8}$	crc8	crc7	crc6	crc5	crc4	crc3	crc2	crc1	crc0
CRC1	$\overline{D8}$	crc17	crc16	crc15	crc14	crc13	crc12	crc11	crc10	crc9

5. 应用

可以用一个或多个串行接口在设备之间传输各种视频数据。

❑ RGBK-隔行SDTV

BT.799和BT.1303也定义了R'G'B'K串行接口。图6-10中的两个10位R'G'B'K并行数据流使用270或360Mbit/s串行接口进行串行化。

❑ 4：2：2 YCbCr-逐行SDTV

ITU-R BT.1362和SMPTE 294M也定义了4：2：2 YCbCr串行接口。图6-12中的这两个10位4：2：2 YCbCr并行数据流用两个270Mbit/s的串行接口进行串行化。

❑ 4：2：2 YCbCr-隔行HDTV

BT.1120和SMPTE 292M也定义了一个4：2：2 YCbCr串行接口。图6-13中的两个10位4：2：2 YCbCr并行数据流首先通过复用的方式组合在一起，然后使用1.485Gbit/s或1.4835Gbit/s的串行接口进行串行化。

6.2 专业视频复合接口

数字复合视频本质上是复合模拟（M）NTSC或（B、D、G、H、I）PAL视频信号的数字版本。采样时钟频率为4倍F_{SC}：大约为14.32MHz［（M）NTSC］和17.73MHz［（B、D、G、H、I）

PAL]。

通常，同时可以支持8位和10位接口，对于10位接口，其中2位用于传输视频数据的小数部分，这可以最小化累积处理误差，同时可以支持10位辅助数据。

表6-9列出了数字复合信号的电平。视频数据可以不使用10位值0x000~0x003和0x3FC~0x3FF，或8位值0x00和0xFF，因为这些值要用于传输时序信息。

表6-9　数字复合视频信号的10位视频量化等级

视频量化等级	(M) NTSC	(B、D、G、H、I) PAL
峰值色度	972	1040（限于1023内）
白色	800	844
峰值色同步	352	380
黑色	280	256
消隐	240	256
峰值脉冲	128	128
峰值色度	104	128
同步	16	4

6.2.1　NTSC视频时序

每个扫描线总共有910个采样点，如图6-22所示。水平采样点计数0对应于有效视频数据的起点，水平计数768对应于水平消隐的开始。

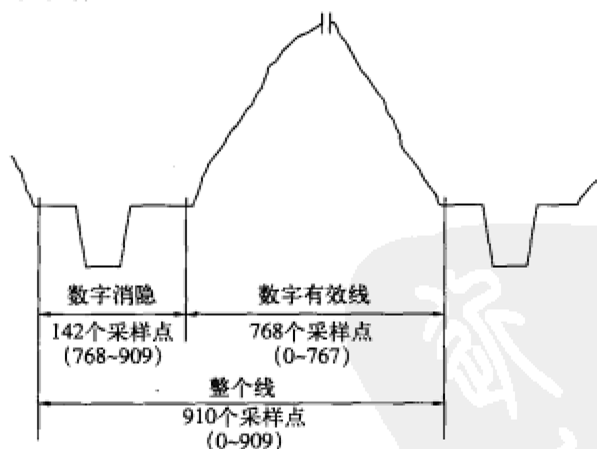


图6-22　数字复合（M）NTSC模拟和数字时序关系

采样是沿着±I和±Q轴（33°、123°、213°和303°）进行的。采样相位位于线10的水平计数0处，场1在+I轴上（123°）。

图6-23和表6-10至表6-12定义了同步边值和始于同步边值的水平计数。一个色同步周期（burst cycle）具有8位值：45、83、75和37。色同步包络（burst envelope）开始于水平计数857处，持续43个时钟周期，如表6-10所示。注意，并没有对脉冲的峰值幅度进行采样。

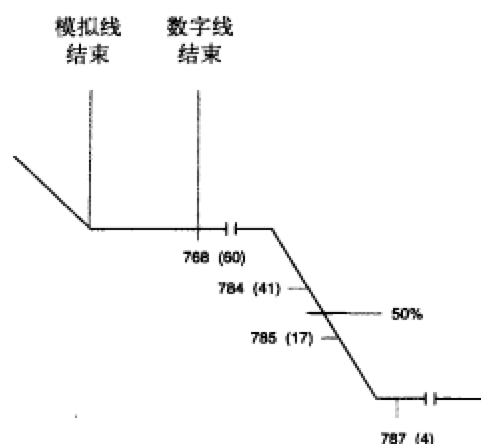


图6-23 数字复合（M）NTSC同步时序。括号中为与8位采样值对应的水平计数

表6-10 在水平消隐期内数字复合（M）NTSC视频数据的数字值

采样点	8位十六进制值		10位十六进制值	
	场1和3	场2和4	场1和3	场2和4
768~782	3C	3C	0F0	0F0
783	3A	3A	0E9	0E9
784	29	29	0A4	0A4
785	11	11	044	044
786	04	04	011	011
787~849	04	04	010	010
850	06	06	017	017
851	17	17	05C	05C
852	2F	2F	0BC	0BC
853	3C	3C	0EF	0EF
854~856	3C	3C	0F0	0F0
857	3C	3C	0F0	0F0
858	3D	3B	0F4	0EC
859	37	41	0DC	104
860	36	42	0D6	10A
861	4B	2D	12C	0B4
862	49	2F	123	0BD
863	25	53	096	14A
864	2D	4B	0B3	12D
865	53	25	14E	092
866	4B	2D	12D	0B3
867	25	53	092	14E
868	2D	4B	0B3	12D
869	53	25	14E	092
870	4B	2D	12D	0B3
871	25	53	092	14E
872	2D	4B	0B3	12D

（续）

采样点	8位十六进制值		10位十六进制值	
	场1和3	场2和4	场1和3	场2和4
873	53	25	14E	092
874	4B	2D	12D	0B3
875	25	53	092	14E
876	2D	4B	0B3	12D
877	53	25	14E	092
878	4B	2D	12D	0B3
879	25	53	092	14E
880	2D	4B	0B3	12D
881	53	25	14E	092
882	4B	2D	12D	0B3
883	25	53	092	14E
884	2D	4B	0B3	12D
885	53	25	14E	092
886	4B	2D	12D	0B3
887	25	53	092	14E
888	2D	4B	0B3	12D
889	53	25	14E	092
890	4B	2D	12D	0B3
891	25	53	092	14E
892	2D	4B	0B3	12D
893	53	25	14E	092
894	4A	2E	129	0B7
895	2A	4E	0A6	13A
896	33	45	0CD	113
897	44	34	112	0CE
898	3F	39	0FA	0E6
899	3B	3D	0EC	0F4
900~909	3C	3C	0F0	0F0

表6-11 数字复合（M）NTSC视频信号场消隐期内的均衡色同步值

场1、3采样点	8位十六进制值	10位十六进制值	场2、4采样点	8位十六进制值	10位十六进制值
768~782	3C	0F0	313~327	3C	0F0
783	3A	0E9	328	3A	0E9
784	29	0A4	329	29	0A4
785	11	044	330	11	044
786	04	011	331	04	011
787~815	04	010	332~360	04	010
816	06	017	361	06	017
817	17	05C	362	17	05C
818	2F	0BC	363	2F	0BC
819	3C	0EF	364	3C	0EF

（续）

场1、3采样点	8位十六进制值	10位十六进制值	场2、4采样点	8位十六进制值	10位十六进制值
820~327	3C	0F0	365~782	3C	0F0
328	3A	0E9	783	3A	0E9
329	29	0A4	784	29	0A4
330	11	044	785	11	044
331	04	011	786	04	011
332~360	04	010	787~815	04	010
361	06	017	816	06	017
362	17	05C	817	17	05C
363	2F	0BC	818	2F	0BC
364	3C	0EF	819	3C	0EF
365~782	3C	0F0	820~327	3C	0F0

表6-12 数字复合（M）NTSC视频信号场消隐期内的锯齿脉冲值

场1、3			场2、4		
采样点	8位十六进制值	10位十六进制值	采样点	8位十六进制值	10位十六进制值
782	3C	0F0	327	3C	0F0
783	3A	0E9	328	3A	0E9
784	29	0A4	329	29	0A4
785	11	044	330	11	044
786	04	011	331	04	011
787~260	04	010	332~715	04	010
261	06	017	716	06	017
262	17	05C	717	17	05C
263	2F	0BC	718	2F	0BC
264	3C	0EF	719	3C	0EF
265~327	3C	0F0	720~782	3C	0F0
328	3A	0E9	783	3A	0E9
329	29	0A4	784	29	0A4
330	11	044	785	11	044
331	04	011	786	04	011
332~715	04	010	787~260	04	010
716	06	017	261	06	017
717	17	05C	262	17	05C
718	2F	0BC	263	2F	0BC
719	3C	0EF	264	3C	0EF
720~782	3C	0F0	265~327	3C	0F0

为了保持0SCH相位，水平计数784开始于线同步下降沿50%之前25.6ns处，而水平计数785在线同步下降沿50%之前44.2ns开始（副载波相位为57°处）。

129

6.2.2 PAL视频时序

除了每帧中有两线是1137个采样点外，其余每线总共有1135个采样点，因此每帧总共有

709 379个采样点。图6-24表示的是典型的线时序。水平计数0对应于有效视频的起点，水平计数948对应于行消隐的开始。

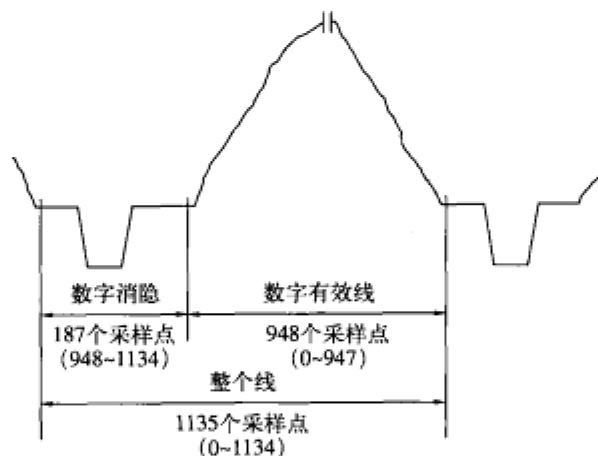


图6-24 数字复合（B、D、G、H、I）PAL模拟和数字时序关系

采样是沿着±U和±V轴（0°、90°、180°和270°）进行的，采样相位位于+V轴（90°）上场1的线1的水平计数0处。

8位色同步值为95、64、32和64，并不断地重复。摇摆色同步（swinging burst）导致峰值脉冲（32和95）和0脉冲（64）采样点的位置发生改变。色同步包络开始于水平计数1058，持续40个时钟周期。

与（M）NTSC的情况不同，在PAL中，采样不是H-凝聚环的（H-coherent），因此不同的扫描线之间，同步脉冲的位置是不同的。交替色同步（alternate burst）采样点值为64时的相位定义为0SCH相位。

6.2.3 辅助数据

辅助数据包用于在消隐期内传输信息（如数字音频、字幕和图文数据）。ITU-R BT.1364和SMPTE 291M描述了辅助数据的格式。

辅助数据格式与本章前面讨论的数字分量视频是一样的。然而，没有使用3个字节的前导，而是一个字的辅助数据标志，具有10位的$3FC_H$值。在TRS-ID后也可能有多个辅助数据标志，每个标志都给出了另一个辅助数据包的开始。

辅助数据可以在下面的字编号边界（见图6-25至图6-30）内提供。

NTSC	PAL	
795~849	972~1035	线同步周期
795~815	972~994	均衡色同步周期
340~360	404~426	
795~260	972~302	场同步周期
340~715	404~869	

130 ~ 135

用户数据可能不使用10位值0x000~0x003和0x3FC~0x3FF，或8位值0x00和0xFF，因为这些值要用于传输时序信息。

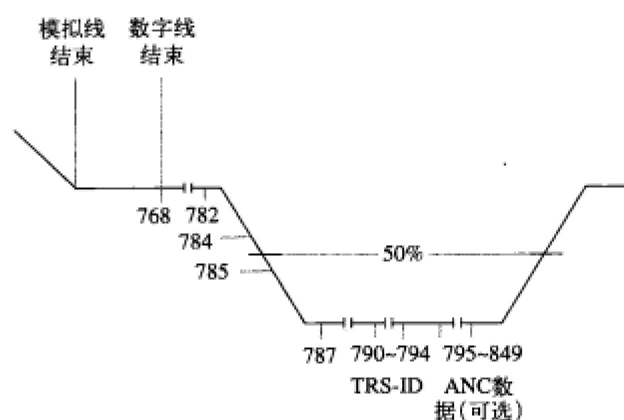


图6-25 在线同步期内（M）NTSC TRS-ID和辅助数据的位置

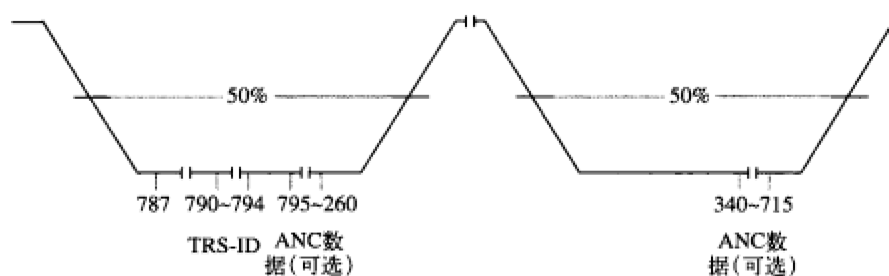


图6-26 在场同步期内（M）NTSC TRS-ID和辅助数据的位置

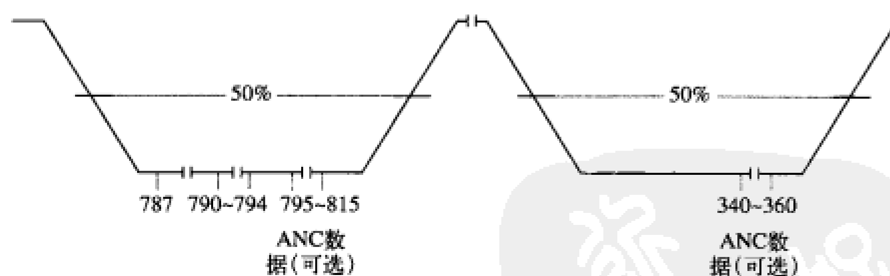


图6-27 在均衡化色同步期内进行（M）NTSC TRS-ID和辅助数据定位

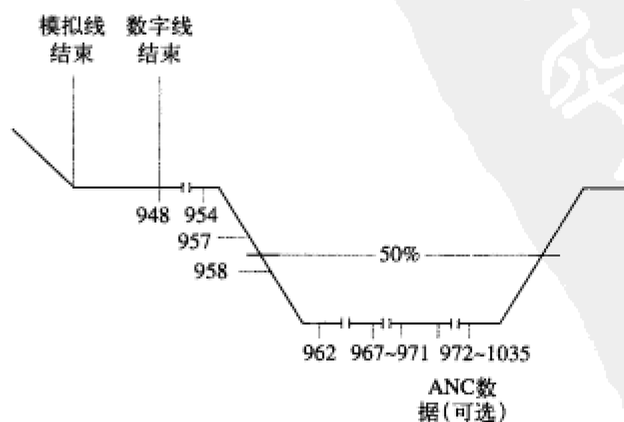


图6-28 在线同步期内进行（B、D、G、H和I）PAL TRS-ID和辅助数据定位

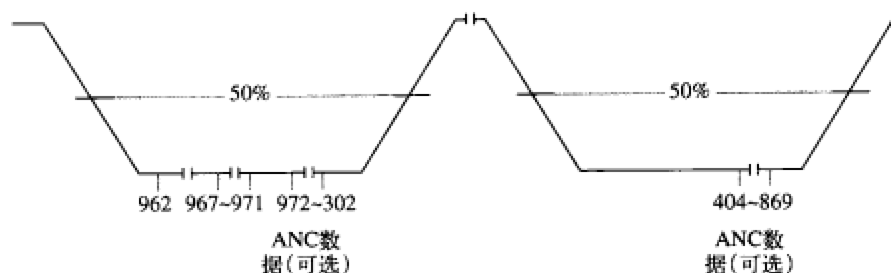


图6-29 在场同步期内进行（B、D、G、H和I）PAL TRS-ID和辅助数据定位

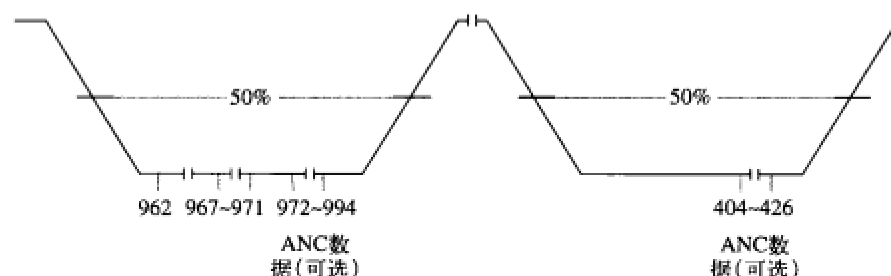


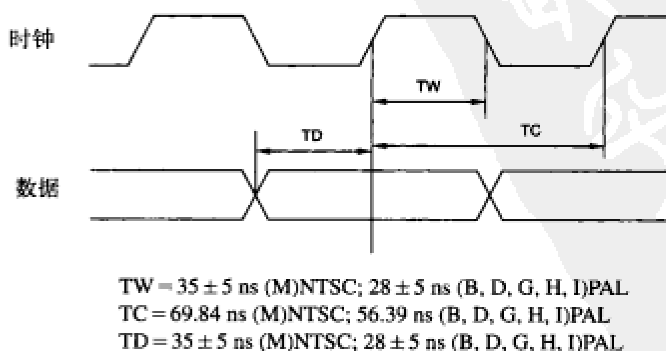
图6-30 在均衡色同步期内进行（B、D、G、H和I）PAL TRS-ID和辅助数据定位

6.2.4 并行接口

这种接口是一种SMPTE 244M的25引脚并行接口，它是基于另一种用于传输27MHz 4：2：2数字视频（见表6-5）接口的，除了时序存在差异外，二者都是相同的。这个接口用于传输SDTV分辨率的数字复合数据。传输的是8位或10位数据和4倍F_{SC}时钟信息。

信号电平与ECL兼容的平衡驱动设备和接收设备是兼容的。生成器必须具有平衡输出，其最大源阻抗为110Ω，信号在110Ω负载上压降的峰间测量值必须为0.8~2.0V。在接收端，传输线必须终止于110±10Ω的负载。

时钟信号为4倍F_{SC}方波，时钟脉冲宽度为35±5ns[（M）NTSC]或28±5ns[（B、D、G、H、I）PAL]。时钟信号的正跃迁发生在数据跃迁的中点，数据跃迁的误差容限为±5ns（如图6-31所示）。



TW = 35 ± 5 ns (M)NTSC; 28 ± 5 ns (B, D, G, H, I)PAL
TC = 69.84 ns (M)NTSC; 56.39 ns (B, D, G, H, I)PAL
TD = 35 ± 5 ns (M)NTSC; 28 ± 5 ns (B, D, G, H, I)PAL

图6-31 数字复合视频并行接口波形

为了可以在50~200m的连接长度上进行可靠的操作，接收端必须使用频率均衡，频率均衡的典型特征见图6-3。这个例子中，有线连接在直到长度为0的各种长度内都可以进行可靠的操作。

6.2.5 串行接口

并行格式可以转换为SMPTE 259M串行格式（图6-32），准许数据通过75Ω的同轴电缆（或光纤）进行传输。这类接口将14.32MHz或17.73MHz的并行数据流转换为143Mbit/s或177Mbit/s串行数据流。利用14.32MHz或17.73MHz的时钟信号，10倍PLL可以生成143MHz或177MHz的时钟。

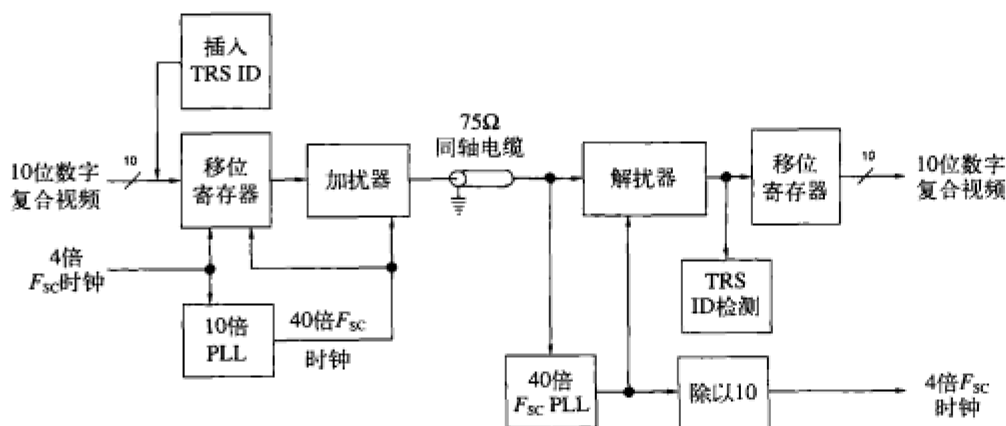


图6-32 串行接口框图

对于有线连接，生成器具有不平衡的输出，其源阻抗为75Ω，信号在75Ω负载山测得的峰间电压必须为0.8±10%。接收端的输入阻抗为75Ω。

10位数据经过串行化（LSB优先）及加扰和非极化NRZI算法处理：

$$G(x)=(x^9+x^4+1)(x+1)$$

该算法与前面讨论过的数字分量视频所使用的算法是一样的。在8位环境下，串行化之前，8位数据需要在最低数据位上附加2位0。

输入到加扰器（见图6-20）中的信号使用正逻辑（最高电压表示逻辑1，最低电压表示逻辑0）。格式化的串行数据是以40倍F_{SC}频率输出的。

在接收端，锁相同步是通过检测TRS-ID序列来完成的。针对每行扫描线，不断地对PLL进行微调以确保能够检测到这些模式，避免发生位跳。恢复的10倍时钟需要除以10来生成4倍的F_{SC}采样时钟。串行数据进行高低频率均衡，并进行解扰（见图6-21）和解串行化。

137 ~ 139

6.2.6 TRS-ID

在使用串行接口时，在线同步时间内，必须在数字视频数据流中插入一个特定的由5字组成的序列，即TRS-ID。TRS-ID仅出现在紧跟同步前沿处，同步前沿指明了水平跃迁，并占据水平计数790~794（包含于NTSC）或967~971（包括于PAL）。表6-13给出了TRS-ID的数据格式，图6-25到图6-30表示的是数字复合（M）NTSC和（B、D、G、H、I）PAL视频信号中TRS-ID的位置。

表6-13 TRS-ID格式

	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
TRS字0	1	1	1	1	1	1	1	1	1	1
TRS字	0	0	0	0	0	0	0	0	0	0
TRS字2	0	0	0	0	0	0	0	0	0	0
TRS字3	0	0	0	0	0	0	0	0	0	0
线编号ID	$\overline{D8}$	EP	线编号ID							

注：EP＝D0~D7的偶校验。

在水平计数794线（NTSC）或971线（PAL）处，编号ID字的定义见表6-14。

表6-14　水平计数794（NTSC）或971（PAL）处的线编号ID字

D2	D1	D0	（M）NTSC	（B、D、G、H、I）PAL
0	0	0	场1中的1~263线	场1中的1~313线
0	0	1	场2中的264~525线	场2中的314~625线
0	1	0	场3中的1~263线	场3中的1~313线
0	1	1	场4中的264~525线	场4中的314~625线
1	0	0	未使用	场5中的1~313线
1	0	1	未使用	场6中的314~625线
1	1	0	未使用	场7中的1~313线
1	1	1	未使用	场8中的314~625线

D7~D3	（M）NTSC	（B、D、G、H、I）PAL
$1 \leq x \leq 30$	线编号1~30[264~293]	线编号1~30[314~343]
$x=31$	线编号≥31[294]	线编号≥31[344]
$x=0$	未使用	未使用

由于具有25Hz的偏移，PAL要求相对于水平同步对TRS-ID的位置进行复位，并且要求每个场都要进行一次复位，复位的位置只能在625~4和313~317线中的某一线上进行。除了用于复位的两线有1137个采样点外，其余所有的线都是1135个采样点。这两个额外采样点的编号为1135和1136，它们的位置在第一个有效图像采样点（采样点0）之前。

由于具有这25Hz的偏移，对每线的采样都要稍微提前。TRS-ID初始位置的确定应该在场1的线1上，或附近的线上完成。TRS-ID的位置总是开始于采样点967处，但由于25Hz的偏移，离同步边沿的距离总是变化的。

140 ~ 142

6.3　专业视频传输接口

6.3.1　串行数据传输接口

SMPTE 305M和ITU-R BT.1381定义了一个串行数据传输接口（SDTI），该接口能够实现设备之间的数据传输。物理层使用270Mbit/s或360Mbit/s的BT.656、BT.1302和SMPTE 259数字分量视频串行接口。图6-33表示的是这种信号的格式。

图6-33　SDTI信号格式

紧接EAV序列之后插入一个53字的头，该头指定了信号的源端、目的端和数据格式。表6-15给出了该头的内容。

表6-15　SDTI头结构

	10位数据									
	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
辅助数据标识（ADF）	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1

（续）

	10位数据									
	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
数据ID（DID）	$\overline{D8}$	EP	0	1	0	0	0	0	0	0
SDID	$\overline{D8}$	EP	0	0	0	0	0	0	0	1
数据计数（DC）	$\overline{D8}$	EP	0	0	1	0	1	1	1	0
线编号	$\overline{D8}$	EP	L7	L6	L5	L4	L3	L2	L1	L0
	$\overline{D8}$	EP	0	0	0	0	0	0	L9	L8
线编号CRC	$\overline{D8}$	C8	C7	C6	C5	C4	C3	C2	C1	C0
	$\overline{D8}$	C17	C16	C15	C14	C13	C12	C11	C10	C9
编码与AAI	$\overline{D8}$	EP	AAI3	AAI2	AAI1	AAI0	CD3	CD2	CD1	CD0
目标地址	$\overline{D8}$	EP	DA7	DA6	DA5	DA4	DA3	DA2	DA1	DA0
	$\overline{D8}$	EP	DA15	DA14	DA13	DA12	DA11	DA10	DA9	DA8
	⋮									
	$\overline{D8}$	EP	DA127	DA126	DA125	DA124	DA123	DA122	DA121	DA120
源地址	$\overline{D8}$	EP	SA7	SA6	SA5	SA4	SA3	SA2	SA1	SA0
	$\overline{D8}$	EP	SA15	SA14	SA13	SA12	SA11	SA10	SA9	SA8
	⋮									
	$\overline{D8}$	EP	SA127	SA126	SA125	SA124	SA123	SA122	SA121	SA120
块类型	$\overline{D8}$	EP	BL7	BL6	BL5	BL4	BL3	BL2	BL1	BL0
载荷CRC标识	$\overline{D8}$	EP	0	0	0	0	0	0	0	CRCF
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
头CRC	$\overline{D8}$	C8	C7	C6	C5	C4	C3	C2	C1	C0
	$\overline{D8}$	C17	C16	C15	C14	C13	C12	C11	C10	C9
校验和	$\overline{D8}$	数据ID的D0~D8的和直到最后的头CRC字。预设为全0，回车符被忽略								

注：EP＝D0~D7的偶校验。

在BT.1381内给出了载荷数据的定义，它是通过基于应用的其他标准（如SMPTE 326M）来定义的。载荷数据可以包含MPEG-2节目或传输流、DV流等，可以使用8位字加上偶校验和$\overline{D8}$、或9位字加上$\overline{D8}$。

1. 线编号

线编号规定了1~525（480i系统）内或1~625（576i系统）内的一个值。L0为最低有效位。

2. 线编号CRC

线编号CRC应用于整个线编号的数据ID，并针对所有的10位数据。C0为最低有效位。CRC是一个18位的值，初始值设为全1：

$CRC=x^{18}+x^5+x^4+x^1$

3. 编码与AAI

4位编码值（CD3~CD0）指定了载荷（包含在SAV和EAV序列之间的用户数据）的长度：

0000 4：2：2 YCbCr视频数据

0001 1440字的载荷（使用270Mbit/s接口）

0010 1920字的载荷（使用360Mbit/s接口）

1000 143Mbit/s数字复合视频

4位授权地址标识符（AAI）值，AAI3~AAI0，指定了目的端和源端的地址格式：

0000　未指定格式

0001　IPv6地址

4. 目的地址与源地址

目的地址和源地址指定了目的设备和源设备的地址。当所有的地址为0，并且AAI3－AAI0＝0时，代表的是所有的地址（广播地址）。

5. 块类型

块类型值指定了载荷的划分。BL7~BL6表示的是载荷块的结构：

00　不带ECC的固定块大小

01　带ECC的固定块大小

10　未指定

11　可变块大小

BL5~BL0表示的是块大小固定的分块。大小可变的块由BL7~BL0表示，其值为11000001。

143 ECC格式是与特定应用相关的。

6. 载荷CRC标志

CRCF位指示在载荷的结尾是否提供载荷CRC：

0　不提供CRC

1　提供CRC

7. 头CRC

头CRC用于直至最后保留数据字的所有编码和AAI字，针对整个10位数据。C0为最低有效位。它为一个18位的值，初值设为全1：

$CRC=x^{18}+x^5+x^4+x^1$

6.3.2　高数据率串行数据传输接口

SMPTE 384M和ITU-R BT.1577定义了一个高数据率串行传输接口（HD-SDTI），使得可以在设备之间传输数据。物理层（physical layer）使用1.485（或1.485/1.001）Gbit/s的SMPTE 292M数字分量视频串行接口。

图6-34表示的是信号格式。两个数据通道通过复用的方式形成一个单一的HD-SDTI数据流，使得一个74.25（或74.25/1.001）MHz的数据流占据Y数据空间，另外一个74.25（或74.25/1.001）MHz数据流占据CbCr数据空间。

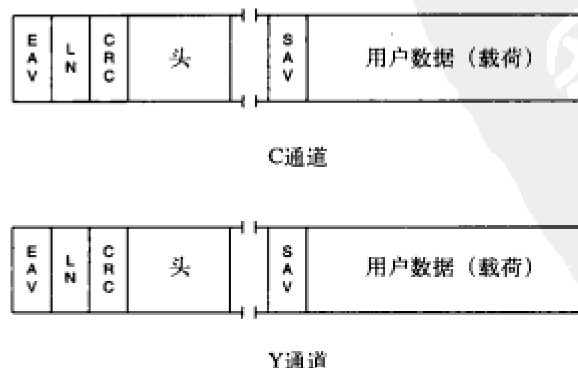


图6-34　HD-SDTI信号格式。LN＝线编号（两个10位的字），CRC＝线编号CRC（两个10位的字）

在线编号CRC数据之后紧接着插入一个49字的头，该头指定了源地址、目的地址和数据格式。表6-16给出了头的内容。

表6-16 HD-SDTI 头格式结构

	10位数据									
	D9（MSB）	D8	D7	D6	D5	D4	D3	D2	D1	D0
辅助数据标识（ADF）	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1
数据ID（DID）	$\overline{D8}$	EP	0	1	0	0	0	0	0	0
SDID	$\overline{D8}$	EP	0	0	0	0	0	0	1	0
数据计数（DC）	$\overline{D8}$	EP	0	0	1	0	1	0	1	0
编码和AAI	$\overline{D8}$	EP	AAI3	AAI2	AAI1	AAI0	CD3	CD2	CD1	CD0
目标地址	$\overline{D8}$	EP	DA7	DA6	DA5	DA4	DA3	DA2	DA1	DA0
	$\overline{D8}$	EP	DA15	DA14	DA13	DA12	DA11	DA10	DA9	DA8
	⋮									
	$\overline{D8}$	EP	SA127	SA126	SA125	SA124	SA123	SA122	SA121	SA120
源地址	$\overline{D8}$	EP	SA7	SA6	SA5	SA4	SA3	SA2	SA1	SA0
	$\overline{D8}$	EP	SA15	SA14	SA13	SA12	SA11	SA10	SA9	SA8
	⋮									
	$\overline{D8}$	EP	SA127	SA126	SA125	SA124	SA123	SA122	SA121	SA120
块类型	$\overline{D8}$	EP	BL7	BL6	BL5	BL4	BL3	BL2	BL1	BL0
载荷CRC标识	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
保留	$\overline{D8}$	EP	0	0	0	0	0	0	0	0
头CRC	$\overline{D8}$	C8	C7	C6	C5	C4	C3	C2	C1	C0
	$\overline{D8}$	C17	C16	C15	C14	C13	C12	C11	C10	C9
校验和	$\overline{D8}$	数据ID的D0~D8的和直到最后的头CRC字。预设为全0，回车符被忽略								

注：EP＝D0~D7的偶校验。

载荷数据是通过其他与应用相关的标准进行定义的。它可能由MPEG-2节目或传输流、DV数据流等组成，使用8位字加上偶校验位和$\overline{D8}$组成或9位字加上$\overline{D8}$组成。

1. 编码和AAI

4位编码值（CD3~CD0）指定了载荷的长度（用户数据包含在SAV和EAV序列之间）：

0000 4：2：2 YCbCr视频数据
0001 1440字的载荷
0010 1920字的载荷
0011 1280字的载荷
1000 143Mbit/s数字复合视频
1001 2304字的载荷（扩展模式）
1010 2400字的载荷（扩展模式）
1011 1440字的载荷（扩展模式）
1100 1728字的载荷（扩展模式）

144~147

1101　2880字的载荷（扩展模式）
1110　3456字的载荷（扩展模式）
1111　3600字的载荷（扩展模式）

扩展模式将SAV序列的时序提前了，缩短了消隐期，因此，载荷数据率保持固定不变的129.6（或129.6/1.001）Mbit/s。

4位授权地址标识（AAI）的格式与SDTI一样。

2. 目的地址和源地址

目的地址和源地址的格式与SDTI是一样的。

3. 块类型

块类型格式与SDTI是一样的。

4. 头CRC

头CRC应用于DID直到最后一个保留数据字，针对整个10位数据。C0为最低有效位。块类型为一个18位的值，其初值设为全1：

148

$CRC=x^{18}+x^5+x^4+x^1$

6.4 IC分量接口

在芯片之间传输数字视频的方案源于专业视频互连标准。典型情况下，针对专业视频应用的芯片都支持每个分量10位或12位的数据。而针对消费类市场的芯片典型情况下每个视频分量使用8位数据。BT.601和BT.656为最流行的接口。

6.4.1 YCbCr值：8位数据

Y的标称取值范围为0x10~0xEB。小于10_H和大于$0xEB_H$的值由于处理的缘故也可以提供。Cb和Cr的标称取值范围为0x10~0xF0。小于0x10和大于0xF0的值由于处理的缘故也可以提供。YCbCr数据可能不使用值00_H和FF_H，因为它们要用于存放时序信息。

在消隐期内，如果没有提供其他数据，那么数据Y的值应该为0x10，数据CbCr 的值应该为0x80。

6.4.2 YCbCr值：10位数据

如果要求具有更高的准确度，典型情况下，专业视频解决方案采用10位的CbCr数据。Y的标称取值范围为0x040~0x3AC。由于处理的需要，小于0x040和大于0x3AC的值也可能被采用。Cb和Cr的标称取值范围为040_H~$3C0_H$。由于处理的需要，小于0x040和大于0x3C0的值也可能被采用。值0x000~0x003和0x3FC~0x3FF可能不会使用，这是为了避免与8位系统的时序内容冲突。

在消隐期内，如果没有提供其他的信息，数据Y的值应该为0x00，数据CbCr的值应该为0x200。

6.4.3 RGB值：8位数据

典型情况下，消费解决方案使用8位R′G′B′数据，表示范围为0x10~0xEB（注意，典型情况下PC使用的范围为0x00~0xFF）。由于处理的缘故，小于0x10和大于0xEB的值也可能要提供。

在消隐期内，如果没有提供其他的信息，R′G′B′数据的应该为0x10。

6.4.4 RGB值：10位数据

为了获得更高的准确度，专业视频标准典型情况下使用10位的R′G′B′数据，标称取值范围为

0x040~0x3AC。由于处理的缘故，也可能需要提供小于0x040和大于0x3AC的值。值0x000~0x003和0x3FC~0x3FF可能不会使用，这是为了避免与8位系统的时序信息冲突。

在消隐期内，如果没有提供额外的数据，那么R'G'B'数据的值应该为0x040。

6.4.5 BT.601视频接口

BT.601视频接口已经使用了多年，它具有控制信号的名称和反映视频标准的时序。所支持的分辨率和采样时钟频率取决于视频标准和宽高比。

设备通常支持多种数据格式，以简化在广泛领域内的使用。

1. 视频数据格式

24位4：4：4 YCbCr数据格式见图6-35。Y、Cb和Cr各有8位，三者以相同的速率进行采样，因此每个采样时钟周期产生24位数据。典型情况下，专业视频解决方案使用30位接口，Y、Cb和Cr数据流每个具有10位。Y0、Cb0和Cr0为最低有效位。

149~150

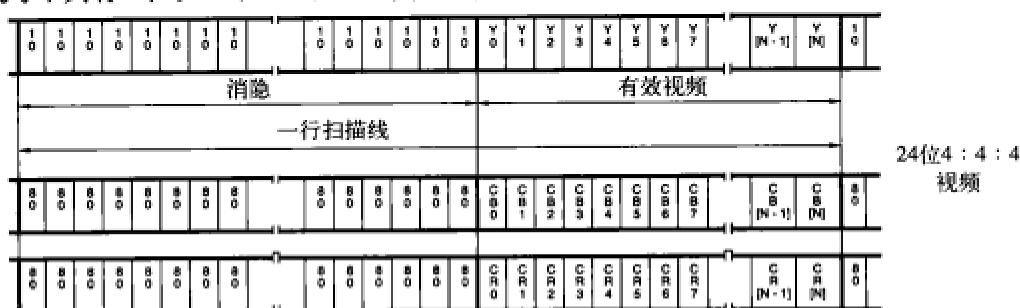


图6-35 24位4：4：4 YCbCr数据格式

16位4：2：2 YCbCr数据格式见图6-36。Cb和Cr以Y的采样率的一半速率进行采样，然后通过复用的方式组合在一起。每个有效字的CbCr数据流总是开始于Cb采样点。典型情况下，专业视频解决方案使用20位接口，Y和CbCr数据流各占10位。

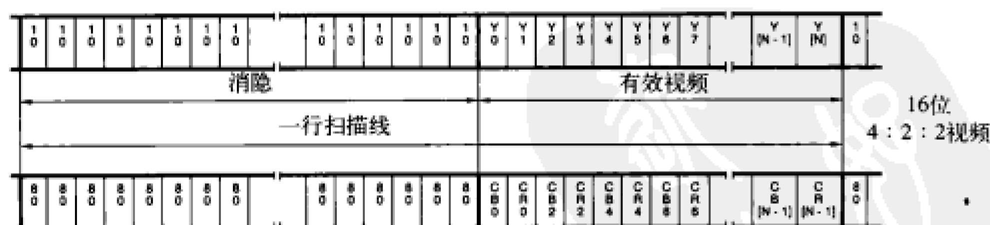


图6-36 16位4：2：2 YCbCr数据格式

8位4：2：2 YCbCr数据格式见图6-37。来自于16位4：2：2 YCbCr数据流的Y和CbCr数据流简单地以2倍采样时钟频率进行复用。有效数据流的YCbCr数据流总是开始于Cb采样点。典型情况下，专业视频解决方案使用10位接口。

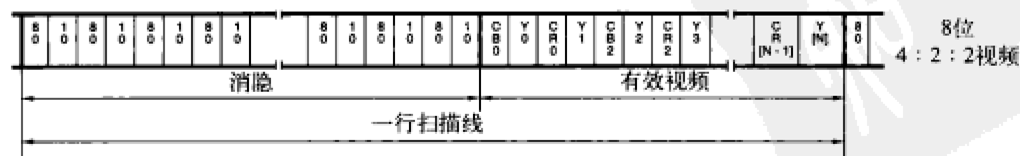


图6-37 8位4：2：2 YCbCr数据格式

表6-17和表6-18表示的是15位RGB、16位的RGB和24位的RGB格式。对于15位的RGB格式，没有使用的位有时被用于alpha信息。R0、G0和B0为最低有效位。

表6-17　在12位、16位和24位接口上传输YCbCr和RGB数据。*很多设计选择使用红色通道来传输复用CbCr数据

24位RGB			16位RGB	15位RGB	24位4：4：4 YCbCr			16位4：2：2	8位4：2：2
单时钟边沿	双时钟边沿		（5、6、5）	（5、5、5）	单时钟边沿	双时钟边沿		YCbCr*	YCbCr
R7					Cr7				
R6					Cr6				
R5					Cr5				
R4					Cr4				
R3					Cr3				
R2					Cr2				
R1					Cr1				
R0					Cr0				
G7			R4	–	Y7			Y7	Cb7，Y7，Cr7
G6			R3	R4	Y6			Y6	Cb6，Y6，Cr6
G5			R2	R3	Y5			Y5	Cb5，Y5，Cr5
G4			R1	R2	Y4			Y4	Cb4，Y4，Cr4
G3	G3	R7	R0	R1	Y3	Y3	Cr7	Y3	Cb3，Y3，Cr3
G2	G2	R6	G5	R0	Y2	Y2	Cr6	Y2	Cb2，Y2，Cr2
G1	G1	R5	G4	G4	Y1	Y1	Cr5	Y1	Cb1，Y1，Cr1
G0	G0	R4	G3	G3	Y0	Y0	Cr4	Y0	Cb0，Y0，Cr0
B7	B7	R3	G2	G2	Cb7	Cb7	Cr3	Cb7，Cr7	
B6	B6	R2	G1	G1	Cb6	Cb6	Cr2	Cb6，Cr6	
B5	B5	R1	G0	G0	Cb5	Cb5	Cr1	Cb5，Cr5	
B4	B4	R0	B4	B4	Cb4	Cb4	Cr0	Cb4，Cr4	
B3	B3	G7	B3	B3	Cb3	Cb3	Y7	Cb3，Cr3	
B2	B2	G6	B2	B2	Cb2	Cb2	Y6	Cb2，Cr2	
B1	B1	G5	B1	B1	Cb1	Cb1	Y5	Cb1，Cr1	
B0	B0	G4	B0	B0	Cb0	Cb0	Y4	Cb0，Cr0	

表6-18　在32位接口上传输YCbCr和RGB数据

24位RGB	16位RGB（5、6、5）	15位RGB（5、5、5）	24位4：4：4 YCbCr	16位4：2：2 YCbCr	8位4：2：2 YCbCr
	R4	–		Y7	
	R3	R4		Y6	
	R2	R3		Y5	
	R1	R2		Y4	
	R0	R1		Y3	
	G5	R0		Y2	
	G4	G4		Y1	
	G3	G3		Y0	
R7	G2	G2	Cr7	Cb7，Cr7	
R6	G1	G1	Cr6	Cb6，Cr6	

（续）

24位RGB	16位RGB（5、6、5）	15位RGB（5、5、5）	24位4：4：4 YCbCr	16位4：2：2 YCbCr	8位4：2：2 YCbCr
R5	G0	G0	Cr5	Cb5，Cr5	
R4	B4	B4	Cr4	Cb4，Cr4	
R3	B3	B3	Cr3	Cb3，Cr3	
R2	B2	B2	Cr2	Cb2，Cr2	
R1	B1	B1	Cr1	Cb1，Cr1	
R0	B0	B0	Cr0	Cb0，Cr0	
G7	R4	–	Y7	Y7	Cb7，Y7，Cr7
G6	R3	R4	Y6	Y6	Cb6，Y6，Cr6
G5	R2	R3	Y5	Y5	Cb5，Y5，Cr5
G4	R1	R2	Y4	Y4	Cb4，Y4，Cr4
G3	R0	R1	Y3	Y3	Cb3，Y3，Cr3
G2	G5	R0	Y2	Y2	Cb2，Y2，Cr2
G1	G4	G4	Y1	Y1	Cb1，Y1，Cr1
G0	G3	G3	Y0	Y0	Cb0，Y0，Cr0
B7	G2	G2	Cb7	Cb7，Cr7	
B6	G1	G1	Cb6	Cb6，Cr6	
B5	G0	G0	Cb5	Cb5，Cr5	
B4	B4	B4	Cb4	Cb4，Cr4	
B3	B3	B3	Cb3	Cb3，Cr3	
B2	B2	B2	Cb2	Cb2，Cr2	
B1	B1	B1	Cb1	Cb1，Cr1	
B0	B0	B0	Cb0	Cb0，Cr0	

2. 控制信号

除了视频数据外，还有4个控制信号：

HSYNC#（或HREF）　　水平同步
VSYNC#（或VREF）　　场同步
BLANK#（或ACTIVE）　　消隐
CLK　　1倍或2倍采样时钟

对于8位和10位4：2：2 YCbCr数据格式，CLK为2倍采样时钟速率。对于其他的数据格式，CLK为1倍采样时钟速率。在信号源端，控制信号和视频数据紧接着CLK的上升沿输出。在信号接收端，控制信号和视频数据在CLK上升沿进行采样。

当BLANK#为否时，提供的是有效的R'G'B'或YCbCr视频数据。

在每行扫描线的水平同步时间内，声明HSYNC#同步，其前沿表示新扫描线的开始。HSYNC#声明的时间量通常与视频标准规定的量是一样的。

在每个场或帧的场同步时间内，声明VSYNC#同步，其前沿表示新的场或帧的开始。VSYNC#声明的扫描线编号通常与视频标准的规定是一样的。

对于隔行视频，如果VSYNC#和HSYNC#的前沿是一致的，那么，该场就为场1。如果VSYNC#的前沿位于中线上，该场就是场2。对于非隔行视频，VSYNC#的前沿表示新帧的开始。图6-38给出了HSYNC#和VSYNC#之间的典型关系。

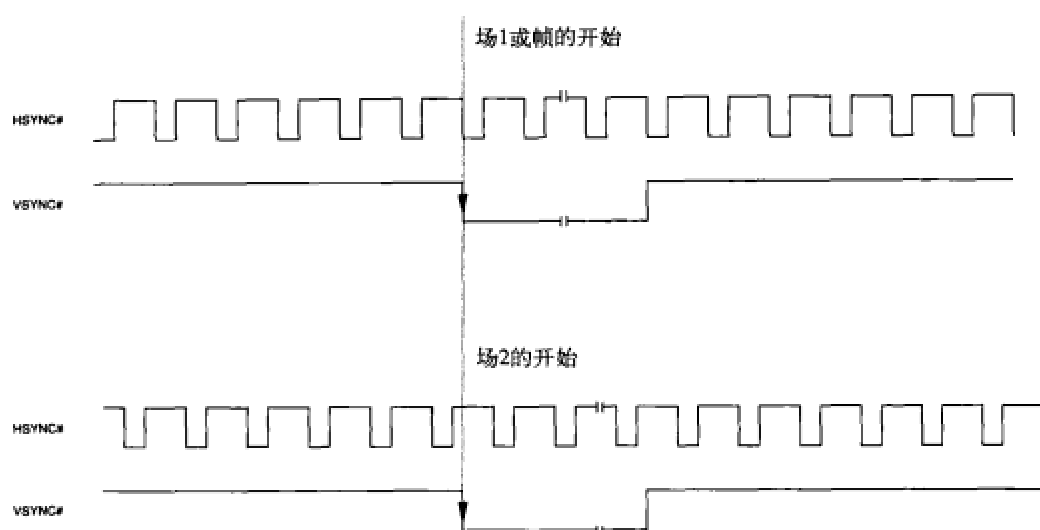


图6-38 HSYNC#和VSYNC#之间的典型关系（没有缩放）。有些产品使用不同的信号名称（如HREF、VREF和ACTIVE），不同的极性和稍微不同的时序

有些产品使用不同的信号名称（如HREF、VREF和ACTIVE）、不同的极性和稍微不同的信号时序。有些产品同时使用时钟的两个边沿也可以传输数据和控制信息，以减少引脚的数量或使得能够处理HDTV的数据率而不需要增加引脚数量。

3. 与接收端相关的注意事项

不应该对每线的采样点数量或行消隐期做出假设。否则，在源端的实现可能不起作用。

为了确保各种信号源之间的兼容，水平计数器应该通过HSYNC#的前沿进行复位，而不是通过BLANK#的后沿来复位。

为了处理现实中的信号源，接收端应该使用一个时间窗来检测提供的是场1还是场2。例如，如果VSYNC#的前沿出现在HSYNC#前沿的1个时钟周期的±64内，那么该场为场1。否则该场为场2。

151
153

有些视频源使用小于0x10的8位数据Y的值来标识同步时序。然而，大多数视频IC并不是这样。另外，为了准许现实世界的视频和测试信号可以以最小的干扰通过，现在许多IC准许数据Y在有效视频期间具有小于0x10的值。这样，接收端在Y通道提供同步时序的设计假定就可能不再成立。

6.4.6 视频模块接口

VMI（视频模块接口）的建立是为了与几个多媒体IC制造商合作。其目的是使得设备之间的视频接口标准化，如MPEG解码器、NTSC/PAL解码器和图形芯片。

1. 视频数据格式

VMI规范规定了8位的4：2：2 YCbCr数据格式，如图6-39所示。很多设备也支持其他的YCbCr和R′G′B′格式，这些已经在6.4.5节讨论过。

2. 控制信号

除了视频数据外，还有4种控制信号：

HREF	水平消隐
VREF	场同步
VACTIVE	有效视频
PIXCLK	2倍采样时钟频率

154



图6-39 一行扫描线的VMI 8位4：2：2 YCbCr数据

对于8位和10位的4：2：2 YCbCr数据格式，PIXCLK为2倍采样时钟频率。对于其他的数据格式，PIXCLK为1倍的采样时钟频率。在源端，控制信号和视频数据紧接着PIXCLK的上升沿输出。在接收端，在PIXCLK的上升沿对控制信号和视频数据进行采样。

当VACTIVE声明为有效时，提供的是有效R'G'B'或YCbCr视频数据。尽管以VACTIVE传输是准许的，但它主要是为了可以利用硬件机制来获取视频数据。对于不支持VACTIVE信号的系统，HREF通常可以以最小的功能代价连入VACTIVE。

为了支持不生成线锁时钟的视频源，也可以使用DVALID#（数据有效）信号。当DVALID#声明为有效时，提供的就是有效数据。

在每行扫描线的有效视频时间内（包括在场消隐期内）声明HREF。

在场同步开始之后经过一半扫描线时，VREF被声明为6倍扫描线时间。

对于隔行视频，VREF的后沿用于对HREF采样。如果HREF声明位有效，那么，该场就为场1，如果HREF声明位无效，那么该场为场2。对于逐行视频，VREF的前沿表示新帧的开始。图6-40表示的是HREF和VREF的典型关系。



图6-40 VMI中HREF与VREF的典型关系（没有缩放）

3. 关于接收端的注意事项

不应该对每线采样点的数量和行消隐期做出假设。否则，针对所有的源时可能会不起作用。

视频数据的输入设置和保持时间分别为5ns和0ns，这是相对于PIXCLK的上升沿的值。

ACTIVE的输入设置和保持时间分别为5ns和0ns，这是相对于PIXCLK的上升沿的值。

155 HREF和VREF的输入设置和保持时间分别为5ns和5ns，这是相对于PIXCLK的上升沿的值。

6.4.7 BT.656接口

用于IC（集成电路）的BT.656接口是基于专业视频的BT.656类并行接口的，已经在本章前面讨论过BT.656类并行接口（图6-1和图6-9）。使用EAV和SAV序列来指示视频时序，减少了所需的引脚数量。用于通用视频格式的H、V和F信号时序见第4章。

BT.656接口使用的是标准的IC信号电平和时序，并支持任意的分辨率。

1. 视频数据格式

BT.656接口使用的是8位或10位4：2：2 YCbCr数据，见图6-1和图6-6。尽管源端可以在EAV和SAV序列中生成4个保护位，但基于芯片之间点到点传输的可靠性，接收端也可以选择忽略它们。

2. 控制信号

CLK为2倍采样时钟频率。在源端，视频数据紧跟CLK的上升沿输出。在接收端，在CLK的上升沿对视频数据进行采样。

为了能够处理HDTV数据传输率，有些设计使用16位或20位YCbCr的接口（本质上是两个156BT.656数据流，一个为数据Y，一个为CbCr数据）或使用时钟的两个边沿传输数据。

6.4.8 缩放视频端口

这是针对笔记本电脑的一个早期标准，缩放视频端口（ZV端口）是PC卡主机适配器和图形控制器之间的点到点单向总线。它使得视频数据可以直接实时地从PC卡传输到图形帧缓存。

PC卡主机适配器具有特定的多媒体模式配置。如果非ZV的PC卡插入扩展槽，主机适配器不会切换到多媒体模式，PC卡以预期的方式工作。一旦ZV卡插入，主适配器将切换到多媒体模式，引脚的功能指派也发生了变化。如表6-19所示，PC卡信号A6~A25、SPKR#、INPACK#和IOIS16#被ZV端口视频信号（Y0~Y7、CbCr0~CbCr7、HREF、VREF和PCLK）以及4通道的音频信号（MCLK、SCLK、LRCK和SDATA）所替换。

表6-19 PC卡与ZV端口信号指派

PC卡信号	ZV端口信号	PC卡信号	ZV端口信号	PC卡信号	ZV端口信号
A25	CbCr7	A17	Y1	A9	Y0
A24	CbCr5	A16	CbCr2	A8	Y2
A23	CbCr3	A15	CbCr4	A7	SCLK
A22	CbCr1	A14	Y6	A6	MCLK
A21	CbCr0	A13	Y4	SPKR#	SDATA
A20	Y7	A12	CbCr6	IOIS16#	PCLK
A19	Y5	A11	VREF	INPACK#	LRCK
A18	Y3	A10	HREF		

1. 视频数据格式

视频数据格式使用的是16位4：2：2 YCbCr数据，如图6-36所示。

2. 控制信号

除了视频数据外，还有3个控制信号：

HREF 水平参考

VREF 场同步

PCLK 1倍采样时钟频率

HREF、VREF和PCLK与本章前面讨论过的VMI接口具有同样的时序。

6.4.9 视频接口端口

VESA VIP（视频接口端口）规范是对前面讨论过的针对IC的BT.656接口的一个增强。尽管这个概念可以很容易地用于其他的应用，但其基本应用就是将最多4种设备接入图形控制芯片。

这类接口有三个部分。

主接口：

VIPCLK	主时钟
HAD0~HAD7	主地址/数据总线
GCTL	主控制

视频接口：

PIXCLK	视频采样时钟
VID0~VID7	低位视频数据总线
VIDA、VIDB	10位数据扩展
XPIXCLK	视频采样时钟
XVID0~XVID7	高位视频数据总线
XVIDA、XVIDB	10位数据扩展

系统接口：

VRST#	复位
VIRQ#	中断请求

主接口信号由图形控制器提供。实际上使用的是PCI接口的2、4或8位的版本。VIPCLK的频率范围为25~30MHz。PIXCLK和XPIXCLK的最高频率分别为75MHz和80MHz。

视频接口

与BT.656接口一样，视频接口需要在8位或10位的4：2：2 YCbCr视频流中插入特定的4字序列，来指示有效视频的开始（SAV）和有效视频的结束（EAV）。这些序列也指示什么时候提供水平和场消隐，以及传输的是哪一个场。

VIP对BT.656 EAV和SAV序列进行了修改，见表6-20。BT.656在状态位中使用了4个保护位（P0~P3），因为BT.656是为设备之间的长缆线连接而设计的。对于芯片到芯片的互连，这种保护是没有必要了，因此，这些数据位就被用于别的用途了。一般的通用视频格式中，H、V和F信号的时序在第4章已经介绍过了。VIP的状态字定义如下：

T＝0，任务B时　　T＝1，任务A时
F＝0，场1时　　F＝1，场2时
V＝1，在垂直消隐期内
H＝0，处于SAV时　　H＝1，处于EAV时

任务位T是可以通过编程来改变的。如果要求与BT.656兼容，那么任务位就应该总是为1。否则，它可能会被用于指示提供的数据流是两个数据流中的哪一个：流A＝1或流B＝0。另外一种替代方式是，当提供2倍过采样VBI数据时，T可以为0，否则T就为1。

表6-20　VIP EAV和SAV序列

	8位数据							
	D7（MSB）	D6	D5	D4	D3	D2	D1	D0
前导帧	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
状态字	T	F	V	H	N	R	0	E

逐行位（N）指示视频源是逐行（1）还是隔行的（0）。

如果当前场为重复场，那么重复位R为1。这种情况只会出现在3：2下拉（pull-down）期。重复位（R）与非重复位（N）是共同起作用的，这使得图形控制器在硬件中处理单场插值和场合并可以与3：2下拉（将在第7章进一步讨论）一样好。

如果在EAV后有另外一个字节的数据，那么额外标志位（extra flag bit）E就为1。表6-21给出了该额外标志字节。仅在EAV期间该位有效。如果在额外字节中的位E为1，那么紧接着就再有一个额外的字节。这样就可以根据需要附加任意多的字节，并通过这种链的方式链接在一起。

158

表6-21　VIP EAV附加字节

	8位数据							
	D7（MSB）	D6	D5	D4	D3	D2	D1	D0
附加字节	$\overline{D0}$	由用户定义						E

与专业视频接口不同，编码0x00可以在有效视频数据期内使用，用它来指示无效的视频采样点。因此这可以用于包容缩放视频和方块像素时序。

❑ 视频数据格式

在8位模式下（见图6-41），除了前面提到的差异外，视频接口与BT.656是类似的。XVID0~XVID7没有使用。

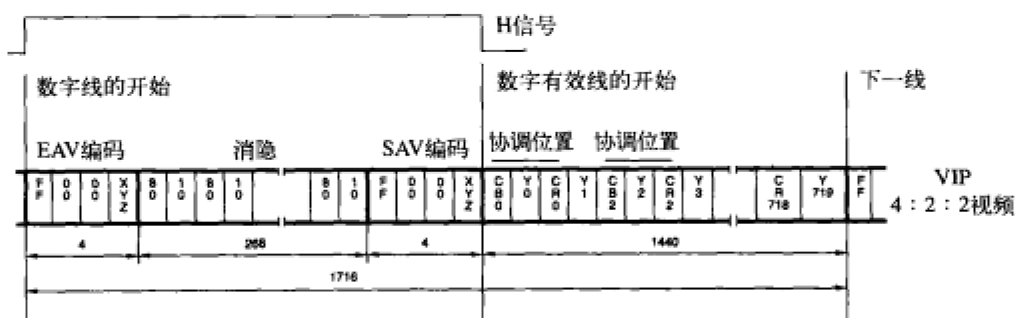


图6-41　VIP 8位接口一行扫描线的数据。480i；每行扫描线有720有效采样点，27 MHz时钟频率

在16位模式下（见图6-42），SAV序列、EAV序列、视频数据Y、辅助数据包头和偶数辅助数据值是通过低8位（VID0~VID7）进行传输的。CbCr视频数据和奇数辅助数据值则是利用高8位（XVID0~XVID7）进行传输的。注意，在有效视频期间，“跳过的数据（skip data）”（值为0x00）也必须出现在16位格式中，这是为了保持16位数据的对齐。

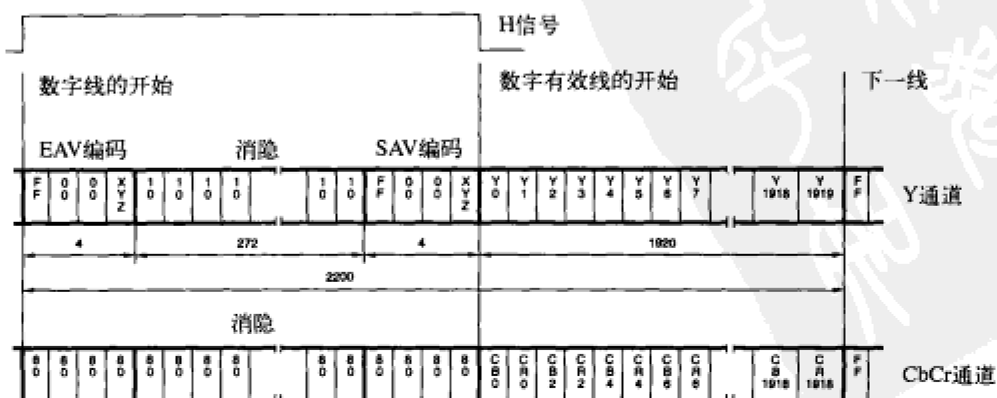


图6-42　VIP 16位接口一行扫描线的数据。1080i；每行扫描线有1920有效采样点，时钟频率为74.176MHz或74.25 MHz

10位视频数据是通过VIDA、VIDB、XVIDA和XVIDB信号来支持的。VIDA和XVIDA为最低有效位。

❑ 辅助数据

辅助数据包用于在消隐期内传输信息（如数字音频、字幕和图文数据），如表6-22所示。与专业视频接口不同，0x00和0xFF这两个值可能被辅助数据所使用。注意，辅助数据格式的定义要比很多专业视频辅助数据格式的定义早，因此可能会存在不匹配的问题。

159
160

表6-22 VIP辅助数据包一般格式

	8位数据							
	D7（MSB）	D6	D5	D4	D3	D2	D1	D0
辅助数据标识（ADF）	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1
数据ID（DID）	$\overline{D6}$	EP	0	1	0	DID2	DID1	DID0
SDID	$\overline{D6}$	EP	用户定义的值					
数据计数（DC）	$\overline{D6}$	EP	DC5	DC4	DC3	DC2	DC1	DC0
内部数据ID0	用户定义的值							
内部数据ID1	用户定义的值							
数据字0	D7	D6	D5	D4	D3	D2	D1	D0
⋮	⋮							
数据字N	D7	D6	D5	D4	D3	D2	D1	D0
校验和	$\overline{D6}$	EP	CS5	CS4	CS3	CS2	CS1	CS0
可选的填充数据	$\overline{D6}$	EP	0	0	0	0	0	0

DID数据域的数据位DID2用于指示提供的是辅助数据场1还是场2：

0 场1

1 场2

DID数据域的数据位DID1~DID0用于指示所提供辅助数据的类型：

00 场的开始

01 分片VBI数据，线1~23

10 场VBI数据的结束， 线23

11 分片VBI数据，线24到场的结束

数据计数值（DC）规定了辅助数据按双字（4字节）计的数目。因此，在DID之后的辅助数据包中的数据字数目必须是4的整倍数。在校验和数据之后可能要加入1~3个可选的占位字节以满足这个要求。

当DID1~DID0为00或10时，就表明没有提供辅助数据或校验和。数据计数值（DC）为00000，并且在包的最后一个场中给出该值。

161

6.5 消费分量接口

在过去几年，开发出了很多在设备之间传输数字视频的解决方案。最初起源于DVI的HDMI是消费类设备最为流行的数字视频接口。

6.5.1 数字可视接口（DVI）

在1998年，数字显示器工作组（Digital Display Working Group，DDWG）组建，目的是解决PC

和VGA显示器之间标准数字视频接口（见图6-43）。DVI 1.0规范于1999年4月正式发布。

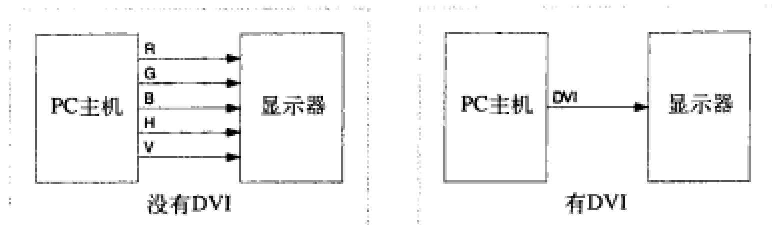


图6-43　用DVI连接VGA显示器和PC主机

DVI是设计用于传输非压缩的实时数字视频的，DVI支持超过1600×1200的PC图形分辨率和包括720p、1080i和1080p的HDTV分辨率。

在2003年，消费类电子企业开始增加了从DVI输出到DVD播放器和有线/卫星机顶盒的支持。DVI输入也开始出现在数字电视和LCD/等离子显示器上了。

1. 技术

DVI是基于数字平板（Digital Flat Panel，DFP）接口的，通过支持更多的格式和时序对其进行了增强。它也包括对高带宽数字内容保护（High-bandwidth Digital Content Protection，HDCP）规范的支持，以抵制视频内容的非授权复制。

DVI也支持VESA的扩展显示标识数据（Extended Display Identification Data，EDID）标准、显示数据通道（Display Data Channel，DDC）标准（用于读取EDID）和显示器时序规范DMT。

DDC和EDID允许自动进行显示检测和配置。扩展显示标识数据（EDID）的建立是为了让显示器具有即插即用的能力。EDID数据存储于显示器中，用于描述所支持的视频格式。EDID数据信息通过DVI提供给源设备，以响应源设备的请求。然后源设备选择其输出格式，同时考虑到原视频流的格式和显示器所支持的格式。源设备负责必要的格式转换以提供给显示器可理解的视频形式。

162

另外，CEA-861标准规定了强制和可选支持的分辨率和时序，还规定了如何包含宽高比和格式信息之类的数据。

2. TMDS链

DVI使用的是最小化传输差分信令（Transition-Minimized Differential Signaling，TMDS）。8位视频数据被转换为10位最小化传输、DC平衡的值，然后再进行串行化。接收端首先对接收到的数据进行反串行化，然后变换回8位数据表示。这样，为了传输数字R′G′B′数据，就需要三个TDMS信号来组成一个TMDS链。

“TFT数据映射”作为所支持的最低要求：每个时钟周期1个像素，每个通道8位，MSB调整。

只可以使用1或2个TMDS链，如图6-44和图6-45所示，这取决于所需的格式和时序。一个支持两个TMDS链的系统必须能够动态地在两种格式之间切换，这两种格式分别为要求单链和要求双链的格式。单个DVI连接头可以处理两个TMDS链。

单TMDS链支持使用25~165MHz的视频采样率的分辨率和时序。使用165~330MHz的视频采样率的分辨率和时序是通过利用两个TMDS链来实现的，每个TMDS链以一半的频率进行操作。这样，两个链就共享相同的时钟，两个链之间也平等地共享带宽。

3. 视频数据格式

典型情况下，24位的R′G′B′数据通过一个链进行传输。对于要求每个颜色分量多于8位的应用，可以采用第二个TMDS链作为附加的最低有效位。

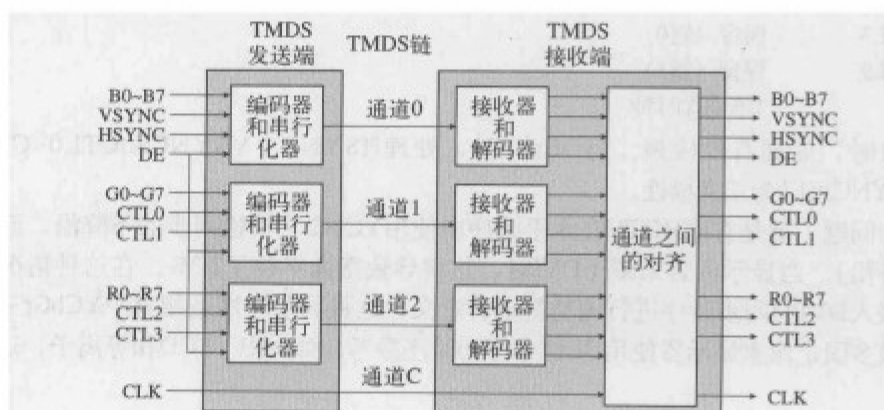


图6-44 DVI单TMDS链

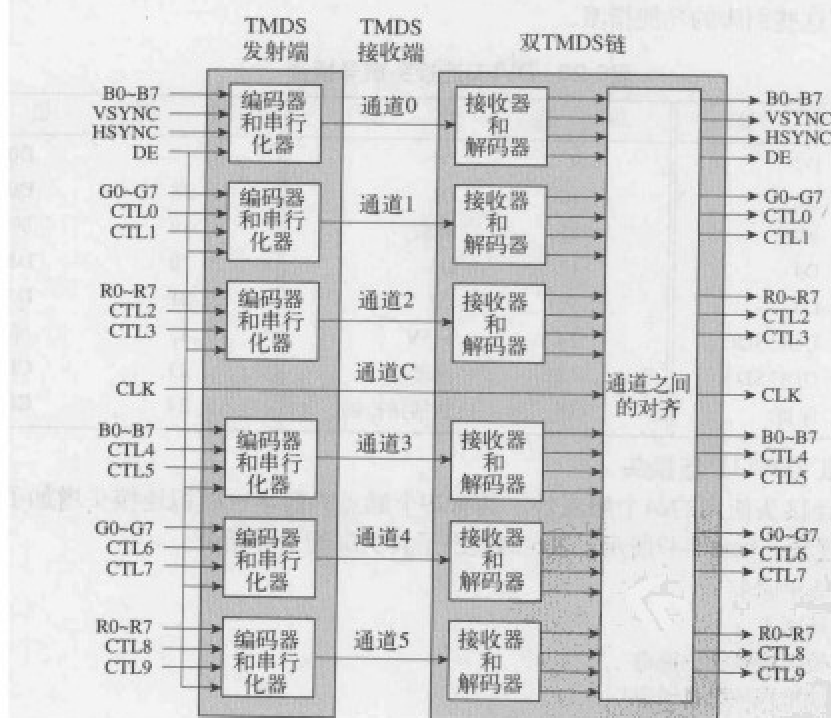


图6-45 DVI双TMDS链

对于PC机的应用，典型情况下，R'G'B'数据的取值范围为0x00~0xFF。对于消费类应用，典型情况下，R'G'B'数据的取值范围为0x10~0xEB（由于处理的需要，偶尔也需要提供小于0x10和大于0xEB的值）。

163
~
164

4. 控制信号

除了视频数据之外，典型情况下，DVI发送和接收芯片在系统中使用多达14个控制信号来实现与其他芯片的连接：

HSYNC	水平同步
VSYNC	场同步
DE	数据使能

CTL0~CTL3　　保留（链0）
CTL4~CTL9　　保留（链1）
CLK　　1倍采样时钟

当DE为1时，处理有效视频。当DE为0时，处理HSYNC、VSYNC和CTL0~CTL9信号。HSYNC和VSYNC可以为任意极性。

存在一个问题，就是有些HDTV在水平时序中使用YCbCr三极性同步的下降沿，而不是使用其中心（上升沿）。当显示内容来源于DVI时，这将导致图像平移了2.3%。在这种情况下，提供一种能力对嵌入DVI的同步时序进行调整就很有意义，这种同步时序是相对于YCbCr三极性同步时序的。有很多固定像素显示器使用DE信号作为时序参考，如DLP、LCD和等离子，这就避免了这个问题。

5. 纯数字（DVI-D）连接头

纯数字连接头（支持双链操作）包含24个触点，它们排列成3行，每行8个触点，如图6-46所示。表6-23列出了这些引脚的功能指派。

表6-23　DVI-D连接头信号指派

引脚	信号	引脚	信号	引脚	信号
1	D2-	9	D1-	17	D0-
2	D2	10	D1	18	D0
3	屏蔽	11	屏蔽	19	屏蔽
4	D4-	12	D3-	20	D5-
5	D4	13	D3	21	D5
6	DDC SCL	14	+5V	22	屏蔽
7	DDC SDA	15	接地	23	CLK
8	保留	16	热插拔检测	24	CLK-

6. 数字-模拟（DVI-I）连接头

除了纯数字连接头使用的24个触点外，具有29个触点的数字-模拟连接头增加了另外5个触点用于支持模拟视频，如图6-47所示。表6-24列出了其引脚的信号指派。

HSYNC　　水平同步
VSYNC　　场同步
RED　　模拟视频红色通道
GREEN　　模拟视频绿色通道
BLUE　　模拟视频蓝色通道

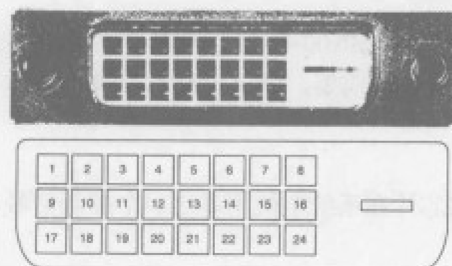


图6-46　DVI-D连接头

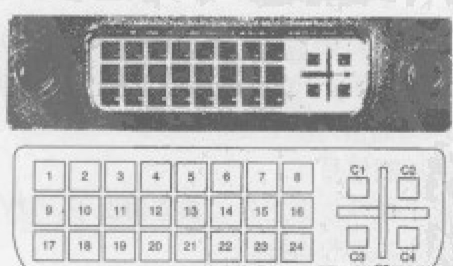


图6-47　DVI-I连接头

表6-24 DVI-I连接头的信号指派

引 脚	信 号	引 脚	信 号	引 脚	信 号
1	D2-	9	D1-	17	D0-
2	D2	10	D1	18	D0
3	屏蔽	11	屏蔽	19	屏蔽
4	D4-	12	D3-	20	D5-
5	D4	13	D3	21	D5
6	DDC SCL	14	+5V	22	屏蔽
7	DDC SDA	15	接地	23	CLK
8	VSYNC	16	热插拔检测	24	CLK-
C1	红色	C2	绿色	C3	蓝色
C4	HSYNC	C5	接地		

模拟信号的操作与标准VGA连接头的情况是一样的。

DVI-A仅作为插入（公头）连接头使用，与其对应的DVI-I连接头的纯模拟引脚相连。DVI-A仅用于适配器有线，因为在适配器中需要与传统的模拟VGA信号相互转换。

165 ~ 166

6.5.2 高清晰度多媒体接口（HDMI）

尽管DVI处理非压缩的实时数字RGB视频到显示器的传输，但是消费类电子企业倾向于更小和更加灵活的解决方案，并且要基于DVI技术。在2002年4月，日立、松下电器、飞利浦、Silicon Image、索尼、汤姆逊和东芝公成立了HDMI工作组。

HDMI能够用一根缆线代替最多8条音频线（7.1个通道）和最多3条视频线，如图6-48所示。在2004年，消费类电子企业开始将HDMI输出加入到DVD播放器和有线/卫星机顶盒中。在2005年，在数字电视和显示器中开始出现HDMI输入。

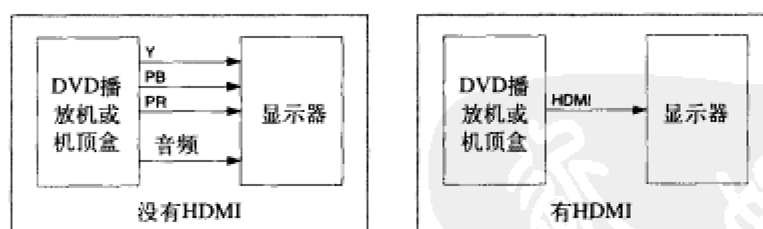


图6-48 利用HDMI消除了消费电子产品中混乱的线缆连接

通过使用适配器线缆，HDMI与使用DVI和CEA-861 DTV档次（profile）的设备是向后兼容的。然而，HDMI先进的特性是不可用的，这些特性包括数字音频、消费电子控制（使得可以在设备之间传输控制指令）和色域元数据等。

1. 技术

基于DVI的HDMI支持VESA的扩展显示标识数据（EDID）标准和显示数据通道（DDC）标准（用于读取EDID）。

除此之外，CEA-861标准还规定了强制和可选支持的分辨率和时序，以及如何包括像宽高比和格式信息这样的数据。

HDMI还支持高频带数字内容保护（HDCP）规范以防止未授权的内容复制。一个常见的问题是源端不能足够频繁地轮流检测TV（要求每秒两次）来查看HDCP线路是否有效。如果当前没有选择TV的HMDI输入，就会出现雪花点，然后就会再次选择HMDI模式。

19引脚A类连接头使用单个TMDS链，所以可以以25~340MHz采样频率传输视频信号。低于25MHz（如480i和576i）的采样速率的视频利用像素复制策略进行传输。

为了支持以高于340MHz的频率对视频信号进行采样，可以使用具有双链能力的29引脚B类连接头。

167

19引脚的C类连接头是为移动应用而设计的，它是A类连接头的一个更小的版本。

2. 视频数据格式

HDMI支持R'G'B'、4：4：4 YCbCr、4：2：2 YcbCr、4：4：4 xvYCC和4：2：2 xvYCC。可以传输每像素的数据位数为24、30、36和48位；深度大于每像素24位的颜色称为"深色（deep color）"。

视频数据的取值范围要么是"全场"（对于8位RGB数据的取值范围为0x00~0xFF），要么是"受限场"（对于8位RGB或Y数据的取值范围为0x10~0xEB，对于8位CbCr数据的取值范围为0x10~0xF0；也可能提供小于或大于这些值的值）。

R'G'B'数据可以是"全场"或"受限场"，但除了640×480分辨率以外，这种分辨率只能是"全场"。

YCbCr和xvYCC视频数据必须总是"受限场"。

3. 音频数据格式

受DVD音频标准的驱动，音频支持包含1~8个非压缩音频流，其采样率可以一直到48kHz、96kHz或192kHz，具体使用哪种采样率取决于视频格式。另外，也可以以192kHz的采样速率传输压缩的多通道音频流。

6.5.3　数字平板接口

开发VESA DFP接口是为了将非压缩数字视频从计算机传输到数字平板显示器。支持VESA即插即用（P&D）标准、扩展显示标识数据（EDID）标准、显示数据通道（DDC）标准和监视器时序规范（DMT）。通过DDC和EDID就可以自动进行显示器检测和配置。仅支持TFT数据映射：每时钟周期1个像素，每通道8位，MSB调整。

和DVI一样，DFP（Digital Flat Panel，数字平板）使用最小传输差分信令（TMDS）。8位视频数据转换为10位最小传输的DC平衡值，然后进行串行化。接收端先将接收到的数据反串行化，然后转换回8位数据。这样，为了传输数字R'G'B'数据需要三个TMDS信号，它们通过复合形成一个TMDS链。线缆长度可以达到5m。

168

1. TMDS链

如图6-49所示，单个TMDS链支持时钟频率要求为22.5~160MHz的格式和时序。

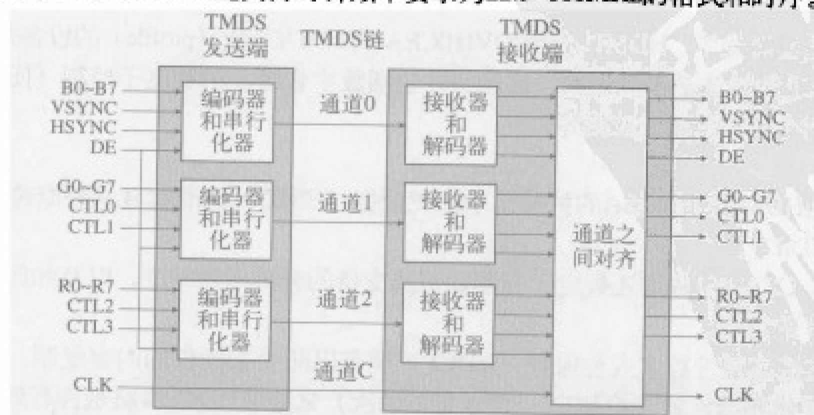


图6-49　DFP TMDS链

2. 视频数据格式

在链上传输24位R'G'B'数据，如图6-49所示。

3. 控制信号

除了视频数据外，典型情况下，DFP发送端和接收端芯片还在系统中使用多达8位控制信号，以实现与其他芯片的连接：

HSYNC　　水平同步
VSYNC　　场同步
DE　　数据使能
CTL0~CTL3　　保留
CLK　　1倍采样时钟

当DE为1时，处理有效视频。当DE为0时，处理HSYNC、VSYNC和CTL0~CTL3信号。HSYNC和CSYNC可以是两种极性的任意一种。

4. 连接头

20引脚迷你D型双边缘（mini-D ribbon，MDR）连接头包含20个触点，排成两行，每行10个，如图6-50所示。表6-25列出了引脚的功能指派。

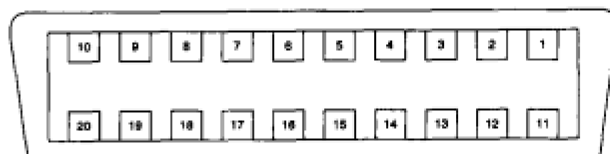


图6-50 DFP连接头

表6-25 DFP连接头信号指派

引脚	信号	引脚	信号
1	D1	11	D2
2	D1-	12	D2-
3	屏蔽	13	屏蔽
4	屏蔽	14	屏蔽
5	CLK	15	D0
6	CLK-	16	D0-
7	接地	17	无连接
8	+5V	18	热插拔检测
9	无连接	19	DDC SDA
10	无连接	20	DDC SCL

169

6.5.4 开放LVDS显示器接口

开发OpenLDI（Open LVDS Display Interface，开放LVDS显示器接口）是为了实现将非压缩数字视频从计算机传输到数字平板显示器。它增强了用于驱动膝上型计算机（笔记本电脑）显示器的FDP链标准，并增加了对VESA的即插即用标准、扩展显示标识数据（EDID）标准和显示数据通道（DDC）标准的支持。DDC和EDID使得可以自动进行显示器的检测和配置。

与DVI和DFP不一样，OpenLDI使用的是低电压差分信令（LVDS）。线缆长度可达10m。

1. LVDS链

如图6-51所示，LVDS链支持时钟频率要求为32.5~160MHz的格式和时序。

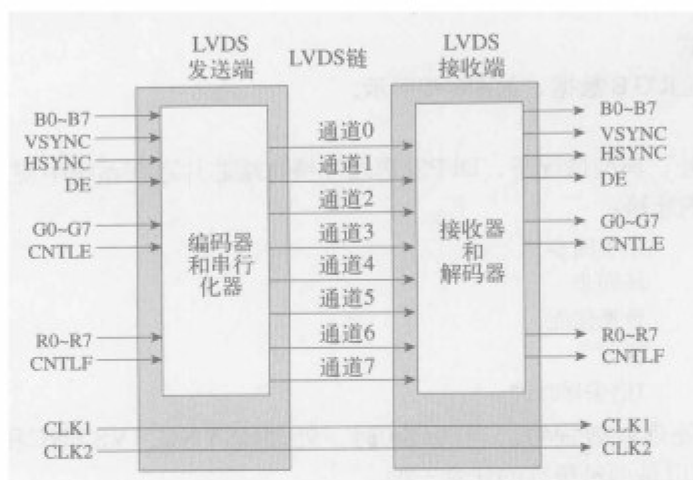


图6-51　OpenLDI LVDS链

LVDS链使用了8根串行数据线（A0~A7）和两个采样时钟线（CLK1和CLK2）。实际使用的串行数据线依赖于像素格式，串行数据率为7倍的采样时钟频率。CLK2信号在双像素模式下用于与FDP链接收端后向兼容。

2. 视频数据格式

在LVDS链上可以传输18位单像素、24位单像素、18位双像素或24位双像素的R'G'B'数据。表6-26给出了像素数据位数与OpenLVI位数的对应关系。

表6-26　OpenLDI 位编号映射

每个像素18位时的位编号	每个像素24位时的编号	OpenLVI位编号
5	7	5
4	6	4
3	5	3
2	4	2
1	3	1
0	2	0
	1	7
	0	6

18位单像素R'G'B'格式使用6位R'G'B'值：R0~R5、G0~G5和B0~B5。OpenLDI串行数据线的A0~A2用于传输数据。

24位单像素R'G'B'格式使用8位R'G'B'值：R0~R7、G0~G7和B0~B7。OpenLDI串行线的A0~A3用于传输数据。

18位双像素R'G'B'格式将两个像素表示为三个6位R'G'B'值高位/低位对：RU0~RU5、GU0~GU5、BU0~BU5、RL0~RL5、GL0~GL5、BL0~BL5。每个高位/低位对表示两个像素。OpenLDI串行数据线的A0~A2和A4~A6用于传输数据。

24位双像素R'G'B'格式将两个像素表示为8位R'G'B'值的3个高位/低位对：RU0~RU7、GU0~GU7、BU0~BU7、RL0~RL7、GL0~GL7、BL0~BL7。每个高位/低位对表示两个像素。OpenLDI串行数据线的A0~A7用于传输数据。

3. 控制信号

除了视频数据外，典型情况下，OpenLDI发送端和接收端芯片在系统中使用最多可达7个控

制信号来实现与其他芯片的连接：

HSYNC 水平同步
VSYNC 场同步
DE 数据使能
CNTLE 保留
CNTLF 保留
CLK1 1倍采样时钟
CLK2 1倍采样时钟

在非平衡操作期间，DE、HSYNC、VSYNC、CNTLE和CNTLF电平作为未编码的位在A2和A6位流中发送。

在平衡操作期间（用于最小化短期和长期DC偏差），在每个A0~A7位流内发送DC平衡位，以指示数据是否被修改或发生过反转。由于剩余的空间用于直接发送控制信号，因此DE电平的发送是通过对CLK1和CLK2信号的下降沿时序进行细微的修改来实现的。HSYNC、VSYNC、CNTLE和CNTLF电平是在消隐期内使用7位编码字分别在A0、A1、A5和A4信号上发送的。

170~171

4. 连接头

除了有每行8个触点的2行引脚外，36引脚迷你D型双边缘（MDR）连接头类似于图6-50所示的连接头。图6-27列出了这些引脚的功能指派。

表6-27 OpenLDI连接头信号指派

引脚	信号	引脚	信号	引脚	信号
1	A0-	13	+5V	25	保留
2	A1-	14	A4-	26	保留
3	A2-	15	A5-	27	接地
4	CLK1-	16	A6-	28	DDC SDA
5	A3-	17	A7-	29	接地
6	接地	18	CLK2-	30	USB-
7	保留	19	A0	31	接地
8	保留	20	A1	32	A4
9	保留	21	A2	33	A5
10	DDC SCL	22	CLK1	34	A6
11	+5V	23	A3	35	A7
12	USB	24	保留	36	CLK2

6.5.5 吉比特视频接口

开发索尼GVIF（Gigabit Video Inter Face，吉比特视频接口）是为了使用单个差分信号传输非压缩数字视频，而不用DVI、DFP和OpenLDI所使用的多个信号进行传输。线缆长度可到达10m。

1. GVIF链

如图6-52所示，GVIF链支持时钟速率要求为20~80MHz的格式和时序。对于要求具有更高时钟频率的应用，可以使用多个GVIF链来实现。

对于8位R′G′B′数据，串行数据速率为24倍采样时钟速率，而对于24位的R′G′B′数据，串行数据采样速率为30倍采样时钟速率。

2. 视频数据格式

可以在GVIF链上传输18位或24位R′G′B′数据和时序信息。18位R′G′B′格式使用三个6位R′G′B′

[172] 值：R0~R5、G0~G5和B0~B5。24位R'G'B'格式使用三个8位R'G'B'值：R0~R7、G0~G7和B0~B7。

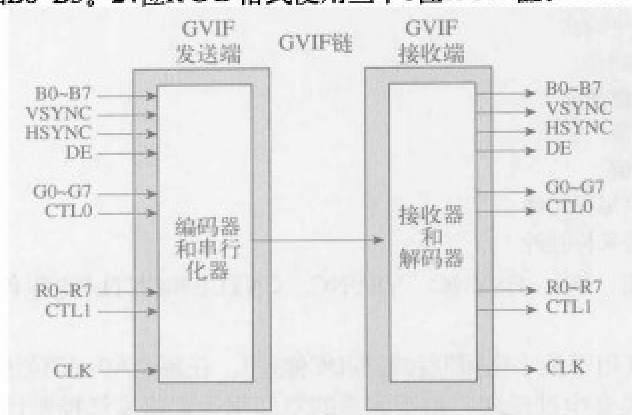


图6-52　GVIF链

为了将18位R'G'B'数据转换为24位数据，首先将R'G'B'数据分割为6个3位的值，然后将这些值转换为6个4位的编码。这就确保了接收端PLL锁具有丰富的转换和良好的DC平衡。

为了将24位R'G'B'数据转换为30位数据，首先需要将R'G'B'数据分割为6个4位的值，然后再将这些值转换为6个5位编码。

3. 控制信号

除了视频数据以外，还有6个控制信号：

HSYNC　　水平同步
VSYNC　　场同步
DE　　数据使能
CTL0　　保留
CTL1　　保留
CLK　　1倍采样时钟

如果HSYNC、VSYNC、DE、CTL0或CTL1中的任何一个信号发生改变，那么在下一个CLK周期内，就会使用一个特定的30位格式。前6位为头数据，它给出了信号HSYNC、VSYNC、DE、CTL0和CTL1的新电平。接着的24位为R'G'B'数据（除了反转的奇数位外都是没有编码的）。

注意，在消隐周期内，可以传输像数字音频这样的非视频数据。CTL信号可以用于指定什么时候提供非视频数据。[173]

6.6　消费传输接口

几种传输接口（如USB 2.0、以太网接口和IEEE 1394）可以用于消费类产品。当然，每个标准都有它们各自的优点和缺点。

6.6.1　USB 2.0

USB 2.0在PC机市场上很著名，主要用于连接PC主机和外部设备，在产品之间利用USB（通用串口总线）2.0传输压缩音频/视频也越来越引起人们的兴趣。

USB 2.0能够处理的传输速率可以达到480Mbit/s，并支持同步模式（isochronous mode）以保证数据传输的时序。这样，USB 2.0就可以很容易将压缩的实时音/视频数据从有线/卫星机顶盒或DVD播放器传输到数字电视。DTCP（数字传输复制保护）可以用于USB上对音频和视频数据内容进行加密。

由于USB的低能耗和广泛应用，许多公司都对使用USB 2.0代替IEEE 1394来实现在产品之间传输压缩音/视频数据感兴趣。然而，有些公司仍然青睐于IEEE 1394，主要是因为该方法对于传输各种类型的数据都有比USB好得多的定义。

移动USB

随着便携设备在大众消费领域的日益流行，人们迫切希望这些设备可以在没有PC机或其他USB宿主机的情况下能够实现相互之间的互连通信。

移动（On-The-Go，OTG）就是要解决这个问题，它通过准许USB设备直接与其他的移动设备进行通信来解决这个问题。同时它也具有更小的USB连接头特征和保持电池使用时间的低能耗特征。

6.6.2 以太网

随着家庭网络、DSL和FTTH（光纤到户）的广泛采用，以太网（Ethernet）已经成为传输数字音频和视频的一个通用接口。起初用于文件传输、实时压缩视频流的有线传输（802.3）或无线传输（802.11）的以太网现在变得越来越普遍了。

以太网支持的传输率可达1Gbit/s。DTCP/IP（用于IP协议的数字传输复制保护）可以用于对有线和无线网络上的音频和视频进行加密。

6.6.3 IEEE 1394

IEEE 1394起初是由苹果公司作为Firewire而建立的。它被设计为设备之间的通用接口，1394规定了其物理特征；独立的、与应用相关的规范描述了如何在1394网络上传输数据。

1394是基于事务（transaction-based）的包交换技术，使用双向串行互连，具有热插拔特性。这使得设备的连接与断开不会影响连接到网络上其他设备的运行。

支持时间敏感数据的确保传输（guaranteed delivery）使得数字音频和视频数据可以实现实时传输。另外，可以加载多个数字音频和视频流。 174

1. 规范

起初，1394-1995规范支持98.304、196.608和393.216Mbit/s的位传输速率。

1394A-2000规范明确了原来模糊的和导致系统互操作问题的一些领域。同时也减少了总线控制、仲裁、总线复位持续时间和组包（concatenation of packets）中的冗余。1394A-2000也引入了先进的节能特性。1394-1995和1394A-2000之间的电子信令方法是相同的，都使用数据过滤（data-strobe，DS）编码和模拟－速度（analog-speed）信号。

1394B-2000规范增加了对位率786.432 Mbit/s、1572.864 Mbit/s和3154.728 Mbit/s的支持，还包括：

- 吉比特以太网中使用的8B/10B编码技术
- 连续双单路复用操作
- 更长的距离（在Cat5上可达100m）
- 将速度信令改为具有更多数字的方法
- 三种类型的端口，Legacy（1395A兼容）、Beta和Bilingual（支持Legacy和Beta）。连接器抠像（keying）确保了在实际中不会发生不兼容的连接。

1394使用了big-endian体系结构，将位0定义为最高有效位。然而，很多处理器都是基于little-endian体系结构的，将位31定义为最高有效位（假定是32位字）。

2. 网络拓扑

和很多网络一样，IEEE 1394没有设计总线主控功能。类似树一样的网络结构具有一个根节

点，在其他设备上分离出逻辑节点（见图6-53）。根节点承担一定的控制功能，根节点是在初始 175 化时选择出来的。一旦选定了根节点，只要它保持加电状态并连入网络，那么它就会一直发挥其功能。

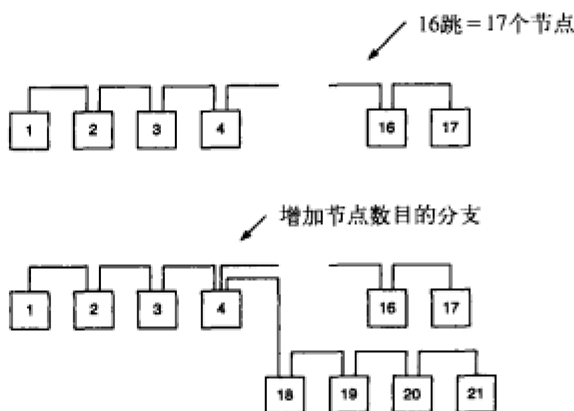


图6-53 IEEE 1394网络拓扑举例

网络可以包含多达63个节点，每个节点（或设备）由一个6位的物理标识编号来指定。多个网络可以通过桥相互连接，系统中最多可以连接1023个网络，每个网络用10位总线ID来表示。二者相结合，系统中16位的地址准许最多有64 449个节点。由于设备地址为64位，其中16位用于指定节点和网络，48位保留用于内存寻址，每个节点准许最多达256TB的内存寻址空间。

3. 节点类型

1394总线上的节点在复杂性和功能上是可变的（按照复杂度从最简单到最复杂的顺序罗列如下）。

事务（transaction）节点负责异步通信，实现了控制状态寄存器（CSR）的最小集，并实现了最小配置ROM。

同步（isochronous）节点增加了24.576MHz时钟，用于对周期时钟寄存器的递增计数，该寄存器是由周期开始包来更新的。

周期主控（cycle master）节点增加了生成8kHz周期起始事件、生成周期起始包的能力，并实现总线定时器寄存器。

同步资源管理器（isochronous resource manager）节点增加了检测坏的自身ID（self-ID）包和确定所选IRM节点ID的能力，并实现了可用通道、可用带宽和总线管理ID寄存器。至少要有一个节点必须能够作为IRM来支持同步通信。

总线管理器（Bus Manager，BM）节点是最复杂的。它可以在拓扑图中存储每个自身ID包，还可以分析该拓扑图以产生整个总线的速度图（speed map）。这两个图用于管理总线。最后，BM必须能够激活周期主控节点，写配置包使得可以对总线进行优化，以及用作电源管理器。

4. 节点端口

在网络拓扑中，单一端口（one-port）设备就是我们所知的“叶（leaf）”设备，因为它处在网络分支的末端。它们可以连接到网络中，但不能对网络进行扩展。

两端口设备可以用于形成顺序链（daisy chained）拓扑。它们既可以连接到网络中，又可以 176 用于对网络的继续扩展，如图6-53所示。具有3个或更多端口的设备能够将网络分支扩展到最多63个分支节点的能力。

注意到在网络中准许无环或并行连接也很重要。没有预留连接头，任何连接头都可以用于将新的设备加入到网络中。

由于1394-1995强制要求任意两个节点之间有线连接的最大跳数为16，因此，如果只有2端口的外围设备可用，那么就可以在网络中包含最多17个外围设备。较新的规范实现了ping包以测试到任意节点的往返跳数延迟，这消除了最多16跳的限制。

对于1394-1995和1394A-2000，使用4引脚或6引脚的连接头。6引脚连接头可以为外围设备提供电源。对于1394B-2002，9引脚Beta和Billigual连接头包含了供电，其中额外的引脚有2个用于信号积分，1个保留供未来使用。

图6-54给出了1394-1995和1394A-2000数据和选通脉冲（strobe）时序。选通脉冲信号改变了数据信号没有改变的每个数据位的状态。

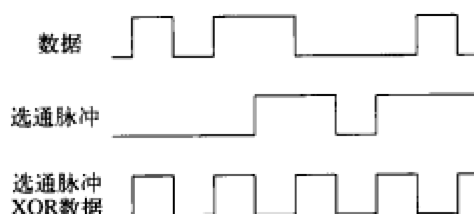


图6-54 IEEE 1394 数据和选通脉冲信号时序

5. 物理层

1394网络的物理硬件拓扑由物理层（PHY）和链路层（LINK）组成，如图6-55所示。1394-1995标准也定义了两个软件层，即传输层和总线管理层，其中有些部分可以在硬件中实现。

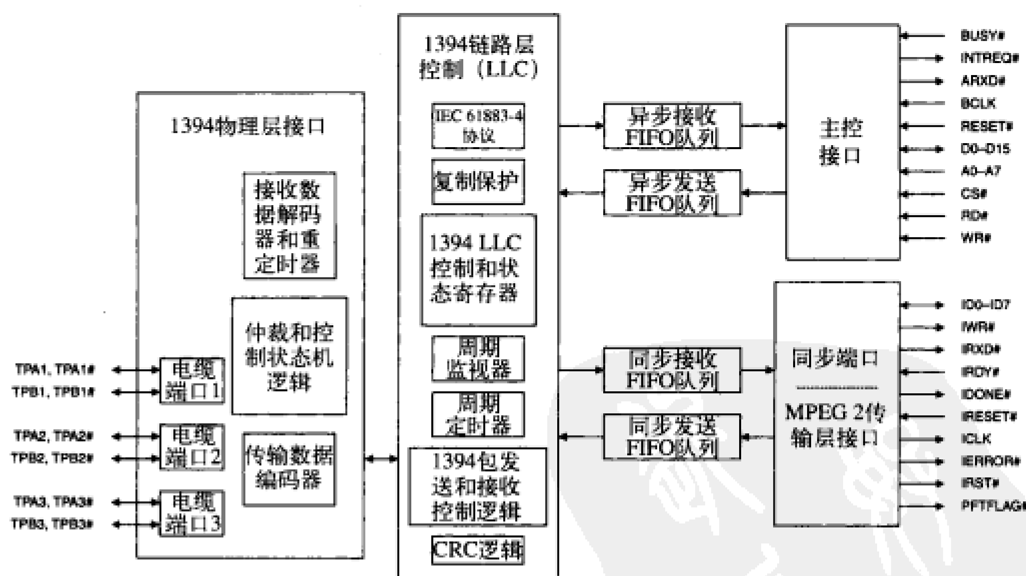


图6-55 IEEE 1394典型的物理和链路层框图

PHY将点到点网络转换为逻辑物理总线。本质上讲，每个节点就是一个数据中继器，因为数据在每个节点重新计时（reclock）。PHY层也定义了到网络的机械与电子连接，还定义了负责开机初始化、仲裁、总线复位感知和数据信号的物理信号电路和逻辑。

6. 链路层

链路层提供了物理层和应用层之间的接口，将数据进行格式化以形成包并在网络上传输。链路层同时支持异步和同步数据。

❑ 异步数据

异步包是确保传输的，因为在异步包被接收到以后，接收端要发送一个确认信息给发送端，如图6-56所示。然而，并不保证带宽。这种类型的通信对于指令、非实时数据和无错误传输的这

类应用是有用的。

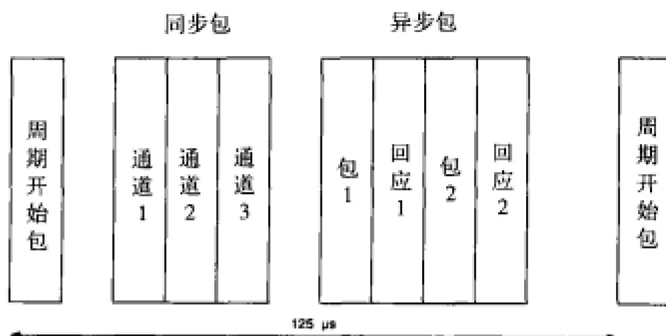


图6-56 IEEE 1394同步和异步包

异步通信包的传输延迟没有保证，这取决于网络通信状态。然而，发送端可以不断地重试，直到接到确认包为止。

异步包的目标是发送到网络上的另一个节点，或发送到网络上的所有节点，但不能对总线上子网的节点进行广播。

最大异步包大小为：

512×（n/100）字节，n＝网络传输速度（单位为Mbit/s）

❑ 同步数据

同步通信的带宽是有保证的，同步通信可以占用网络上的可用带宽可达80%。尽管1394开放主机控制器接口（Open Host Controller Interface，OHCI）当前仅支持4~32个通道，但可用的独立同步通道多达63个。这种类型的通信对于实时音频和视频传输是有用的，因为同步包的最大传输延迟是可计算的，并且可以传输到多个目标。但是，发送端不可以重新尝试发送包。

178

最大的同步包大小为：

1024×（n/100）字节，n＝网络传输速度（单位为Mbit/s）

同步操作可以保证每125μs为一个时间片。由于时间槽是有保证的，并且同步通信的优先级高于异步通信，因此同步带宽是可以保证的。

一旦同步通道建立起来了，那么每个同步周期内在该通道上，发送设备占有所要求的总线时间量就有保证。只有一个设备可以在特定的通道上发送数据，但是任意多的设备可以在一个通道上接收数据。只要总线容量够用，一个设备可以使用多个同步通道。

7. 事务层

事务层支持异步写、读和锁指令。锁结合了一次写和一次读，这种结合是通过发送端和接收端之间产生数据路由的一次往返来实现的，还包括接收端的处理。

8. 总线管理层

总线管理层控制网络在物理、链路和事务层上的功能。

9. 数字传输内容保护（DTCP）

为了阻止未授权的内容复制，建立了DTCP系统。尽管起初是为1394而设计的，但它也适用于任意支持双向通信的数字网络，如USB和以太网。

DTCP支持设备鉴定、内容加密和续保性（设备应该曾经泄密过）。数字传输授权管理机构（Digital Transmission Licensing Administrator，DTLA）许可内容保护系统，并发布加密密钥和设备认证证书。

DTCP概括了内容保护的4个关键因素：

(1) 复制控制信息（CCI）

(2) 认证与密钥交换（AKE）

(3) 内容加密

(4) 系统续保性

❑ 复制控制信息（CCI）

CCI准许内容所有者指定如何使用这些内容，如“不准复制（copy-never）”、“复制一代（copy-one-generation）”、“不准再复制（no-more-copies）”和“自由复制（copy-free）”。DTCP能够安全地在设备之间传输复制一份控制信息。支持两种不同的CCI机制：嵌入CCI（embedded CCI）和加密模式指示器（encryption mode indicator）。

嵌入CCI是将CCI嵌入到内容流中。篡改内容流将导致错误的解码，这就维护了嵌入CCI的完整性。

加密模式指示器（EMI）提供安全的、然而容易访问的CCI传输，这是通过使用同步包头中同步域的两个最高有效位来实现的。设备不需要解码内容就可以立即确定内容流的CCI。如果两个EMI位被篡改过，那么编码和解码模式就不匹配，这将导致不正确的内容解码。 179

❑ 认证与密钥交换（AKE）

在共享内容之前，设备必须首先验证其他的设备是否是可信的。DTCP包括可选的两个认证级别：完全（full）和受限（restricted）。完全认证可以用于受系统保护的所有内容。受限认证仅能保证有效地保护“复制一代”和“复制一份”。

■ 完全认证

兼容设备都由DTLA负责分配一个唯一的公钥/密钥对和一个设备认证证书，二者皆存储于设备内部以防泄露。另外，设备还保存其他一些必要的常量和密钥。

完全验证使用基于数字签名标准（Digital Signature Standard，DSS）的公钥和Diffe-Hellman（DH）秘钥交换算法。DSS是一种用于数字上签名和通过验证数字签名文档的签名来验证数据完整性的一种方法。DH密钥交换用于建立控制－通道对称加密密钥，它准许两个或更多的设备来生成一个共享的密钥。

起初，接收端发送一个请求给源端以交换设备的认证证书和随机查询（random challenge）。然后，每个设备都要计算一个DH密钥交换的第一阶段值。接下来，设备交换签名消息，它们包含以下要点：

(1) 其他设备的随机询问；

(2) DH密钥交换的第一阶段值；

(3) 由设备保存的最新系统续保性消息（SRM）的续保性消息版本号。

设备用其他设备的公钥核对消息签名以验证消息没有被篡改过，同时还要验证其他设备认证证书的完整性。每个设备还要检查嵌入的系统续保性消息（SRM）中的证书吊销列表（Certificate Revocation List，CRL），以验证其他设备中的证书没有由于有损安全而被吊销。如果没有发生错误，两个设备就相互成功地得到验证了，并建立了一个认证密钥。

■ 受限认证

受限认证可以用于源端和接收端之间交换“复制一代”和“复制一份”的内容。它依赖于共享秘密的使用，以对随机询问做出响应。

源端向接收端发起请求，请求其设备ID，并发送一个随机查询。在接收到从源端返回的查询时，接收端计算响应结果并将其发送给源端。

源端将收到的响应与由源端通过使用其服务密钥和接收设备ID生成的类似信息进行比较。如果比较与其自身的计算相匹配，接收端就通过了验证和鉴定。然后源设备和接收设备各自计算一个认证密钥。

❑ 内容加密

为了保证互操作性，所有兼容的设备必须支持56位的M6基准加密。通过使用额外的、可选加密方法可以支持其他的内容保护方式。

❑ 系统续保性

180

支持完全认证的设备可以接收和处理由DTLA创建的、并与内容一起发布的SRM。系统续保性用于保证长期的系统完整性，它通过撤销有害设备的设备ID来实现。

SRM可以有多种方式进行更新：通过其他具有更新列表的兼容设备、具有预先记录好内容的媒体或具有外部通信能力（因特网、电话、通信有线、网络等）的兼容设备。

❑ 操作实例

对于这个例子，让源设备发送一个具有复制保护系统的内容流。

源设备发起内容传输，该内容标记了复制保护状态："复制一代"、"不准复制"、"不准再复制"和"自由复制"。

一旦接收到内容流，接收设备就确定复制保护状态。如果标记为"不准复制"，那么接收设备就请求源设备发起完全认证。如果内容标记为"复制一次"或"不准再复制"（如果支持），那么接收设备将请求完全认证，或如果不支持完全认证，就请求受限认证。

当源设备接收到认证请求后，它就推进所请求的认证类型。如果请求的是完全认证，但源设备只支持受限认证，那么就使用受限认证。

一旦设备完成了认证过程，内容通道加密密钥（内容密钥）就在二者之间进行交换。这个密钥用于在源设备对内容加密，而在接收端对内容进行解密。

10. 1394开放主机控制器接口（OHCI）

1394开放主机控制器接口规范是对1394链路层的实现，并具有一些额外的特性用于支持事务层和总线管理层。它为与1394网络的交互提供了一种标准化的方式。

11. 家庭音视频互操作性（HAVi）

家庭音视频互操作性（HAVi）是1394另一个协议层。HAVi的目标是使得在1394网络中，不管是否有PC主机，1394设备都具有即插即用的互操作性。

12. 串行总线协议（SBP-2）

ANSI串行总线协议2（ANSI Serial Bus Protocol 2，SBP-2）定义了在1394网络上设备之间传输指令和状态包的标准方式，这些设备可以是DVD播放机、打印机、扫描仪、硬盘驱动器和其他设备。

13. IEC 61883规范

特定类型的同步信号，如MPEG-2和IEC 61834、SMPTE 314M以及ITU-R BT.1618数字视频（DV）标准，使用特定的数据传输协议和格式。当数据以同步的方式发送到1394网络上时，要使用特定的打包（packetization）技术。

IEC 61883系列规范定义了在1394网络上传输各种应用相关的数据的具体细节：

IEC 61883-1＝一般规范

IEC 61883-2＝SD-DVCR数据传输，25Mbit/s连续位率

IEC 61883-3＝HD-DVCR数据传输

IEC 61883-4＝MPEG-2 TS数据传输，位率最高可达44Mbit/s

IEC 61883-5＝SDL-DVCR数据传输
IEC 61883-6＝音频和音乐数据传输
IEC 61883-7＝ITU-R BO.1294 B系统的传输

181

❑ IEC 61883-1

IEC 61883-1定义了在1394网络上传输数字音频和视频数据的一般结构。它描述了数字音频和视频的一般包格式，数据流管理和连接管理，也描述了控制指令的一般传输规则。

在同步数据包的数据域起始位置放置了一个公共同步包（Common Isochronous Packet，CIP）头，如图6-57所示。该公共同步包头指定了源节点、数据块大小、数据块计数、时间戳和包含在数据域中的实时数据类型等。

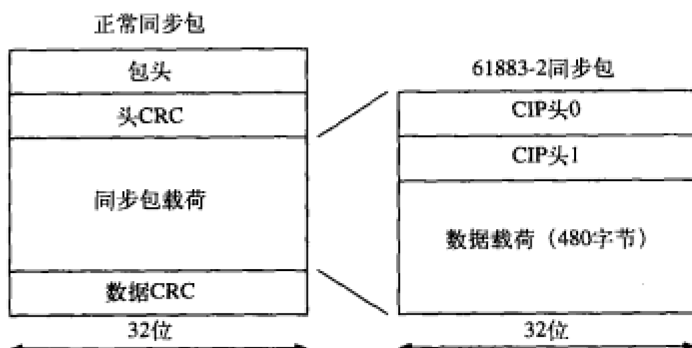


图6-57 61883-2 同步包格式

为了实现设备之间的同步连接，也定义了连接管理过程（Connection Management Procedure，CMP）。

另外，为了实现在1394网络上使用异步数据交换控制指令，定义了功能控制协议（Functional Control Protocol，FCP）。

❑ IEC 61883-2

IEC 61883-2和SMPTE 396M在1394网络上为IEC 61834、SMPTE 314M和ITU-R BT.1618视频数据标准定义了CIP头、数据包格式和传输时序。支持的有效分辨率为720×480（以每秒29.97Hz的帧率）和720×576（以每秒25Hz的帧率）。

DV数据包有488B，由8B的CIP头和480B的DV数据组成，见图6-57。图6-58给出了帧数据结构。

每个720×480的4：1：1 YCbCr帧被压缩成103 950B，因此压缩比为4.9：1。包括冗余和音频数据，数据量增加到了120 000B。

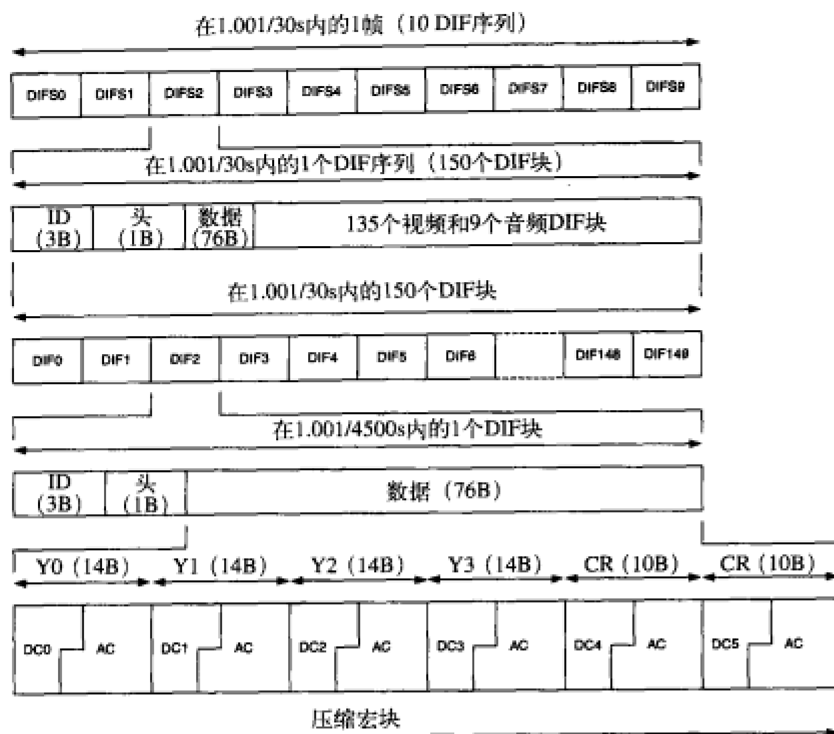
182

压缩的720×480帧被分为10个DIF（Data In Frame）序列。每个DIF序列包含150个DIF块，每个块80B，按如下方式使用：

视频为135个DIF块
音频为9个DIF块
6个DIF块用于头、子码和视频辅助（Video AUXiliary，VAUX）信息

图6-59给出了DIF序列的详细结构。音频DIF块同时包含音频数据和音频辅助数据（Audio AUXiliary，AAUX）。IEC 61834支持4路32kHz、12位非线性音频信号或两路48kHz、44.1kHz或32kHz，16位音频信号。SMPTE 314M和ITU-R BT.1618在25Mbit/s的传输率时支持两路48kHz、16位音频信号，当传输率为50Mbit/s时，可以支持4路48kHz、16位音频信号。视频辅助数据

（VAUX）DIF块包含记录日期和时间、镜头光圈、快门速度、颜色平衡和其他相机设置数据。子码（subcode）DIF块保存各种信息，其中最重要的是时间码。



183

图6-58　720×480系统（4：1：1 YCbCr）中IEC 61834、SMPTE 314M和ITU-R BT.1618包格式

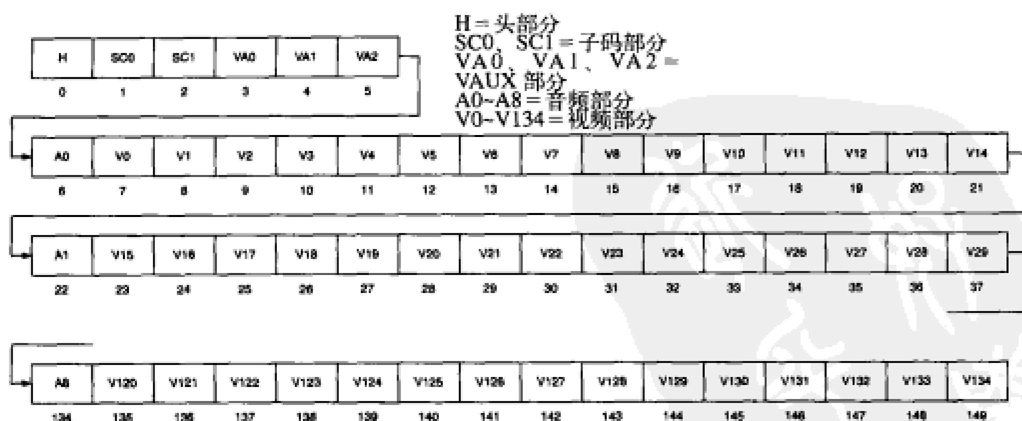


图6-59　IEC 61834、SMPTE 314M和ITU-R BT.1618 DIF序列细节（25Mbit/s）

每个视频DIF块包含80B的压缩宏块数据：

3B的DIF块ID信息

1B的头，包含量化等级数（ANO）和块状态（STA）

14B用于存储Y0、Y1、Y2 和Y3

10B存储Cb和Cr，其中每个字节存储一个Cb和一个Cr

当488B的包通过1394网络时，视频帧的开始就确定了。一旦检测到了帧的开始，那么250个有效数据包就被收集起来并形成一个完整的DV帧，每个包包含6个DIF数据块。每第15个包为一

个空包（null packet），应该丢弃。一旦250个有效数据包存储到了缓冲中，就丢弃CIP头。如果一切正常的话，在帧缓存中就已经有了一个120 000B的压缩DV帧了。

720×576的帧可以使用4：2：0 YCbCr格式（IEC 61834），或4：1：1 YCbCr格式（SMPTE 314M和ITU-R BT.1618）。每个720×576的帧被压缩为124 740 B。包括冗余和音频数据，数据量增加为144 000B，需要300包来传输一个帧。

注意，在1394网络上传输的数据的组织方式与实际的DV记录格式是不同的，因为数字传输不需要错误纠正。另外，尽管图6-59中视频块在序列中有编号，但是序列与视频数据块从左到右、从上到下的传输方式并不是对应的。要打乱（shuffle）压缩的宏块以最小化错误产生的影响，并有助于隐藏错误。视频数据也需要打乱。数据以记录的乱序相同的顺序进行传输。

为了说明视频数据的打乱过程，将DV视频帧组织成50个超块，每个超块由27个宏块组成，如图6-60所示。5个超块为一组（从每个超块栏中取一个），这样一组超块就组成了一个DIF序列。表6-28给出了DIF块的传输顺序。DV数据结构中的其他信息在第11章介绍。

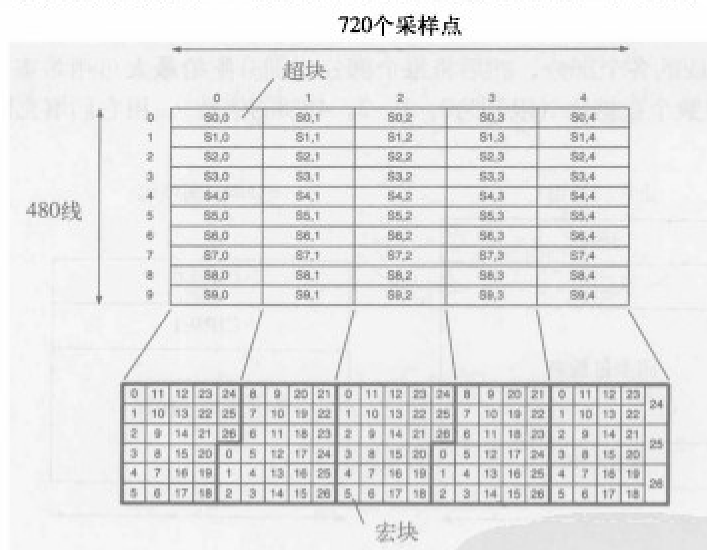


图6-60 超块与宏块的关系（720×480、4：1：1 YCbCr）

表6-28 传输速率为25Mbit/s时的视频DIF块和压缩宏块

DIF序列编号	视频DIF块编号	压缩宏块		DIF序列编号	视频DIF块编号	压缩宏块	
		超块编号	宏块编号			超块编号	宏块编号
0	0	2，2	0	⋮			
	1	6，1	0	$n-1$	0	1，2	0
	2	8，3	0		1	5，1	0
	3	0，0	0		2	7，3	0
	4	4，4	0		3	$n-1$，0	0
	⋮				4	3，4	0
	133	0，0	26		⋮		
	134	4，4	26		133	$n-1$，0	26
1	0	3，2	0		134	3，4	26
	1	7，1	0				
	2	9，3	0				

（续）

DIF序列编号	视频DIF块编号	压缩宏块		DIF序列编号	视频DIF块编号	压缩宏块	
		超块编号	宏块编号			超块编号	宏块编号
1	3	1，0	0				
	4	5，4	0				
		⋮					
	133	1，0	26				
	134	5，4	26				

注：480线系统$n=10$，576线系统$n=12$。

❑ IEC 61883-4

为在1394网络上实现MPEG-2传输流，IEC 61883-4定义了CIP头、数据包格式和传输时间。

184～187

每个同步包大小为192B（188B的MPEG-2数据加上4B的时间戳）的整数倍时效率是最高的，如图6-61所示。然而在很少情况下MPEG数据率是同步数据率的整数倍。这样，将MPEG数据包划分为更小的由24B组成的各个部分，然后将每个部分分别分配给最大可用带宽，这样效率将更高。然后发送端使用整数个数据块（限制为0、1、2、4或8的倍数），用它们填充同步包并增加8B的CIP头。

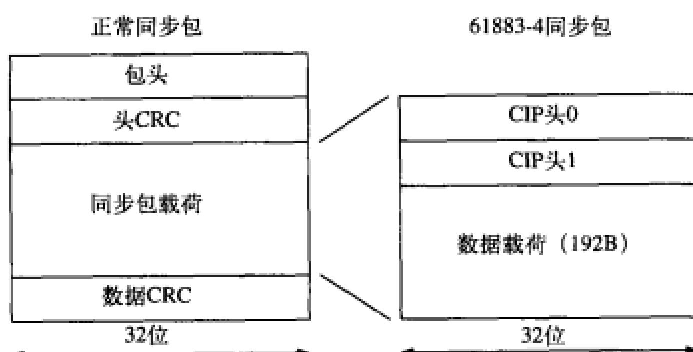


图6-61　61883-4同步包格式

14. 50Mbit/s的DV

与25Mbit/s的DV格式一样，50Mbit/s的DV格式支持720×480i30和720×576i30视频源。然而，50Mbit/s的DV格式使用的是4：2：2 YCbCr，而不是4：1：1 YCbCr。

与前面讨论过的一样，对于25Mbit/s的DV数据流，源包大小为480B（由6个DIF块组成）。250个（576i25系统为300个）大小为480B的包在25Mbit/s的通道上传输。

对于50Mbit/s的DV数据流，源包大小为960B（由12个DIF块组成）。前125个大小为960B的包在一个25Mbit/s的通道上传输，后125个（576i25系统为150个）大小为960B的包在第2个25Mbit/s的通道上传输。

15. 100Mbit/s的DV

100Mbit/s的DV数据流支持1920×1080i30、1920×1080i25和1280×720p60的视频源。1920×1080i30视频源在水平方向上缩放为1280×1080i30。1920×1080i25视频源在垂直方向上缩放为1440×1080i25。1280×720p60的视频源在水平方向上缩放为960×720p60。使用的格式为4：2：2 YCbCr。

100Mbit/s的DV数据流的源包大小为1920B（由24个DIF块组成）。前63个（1080i25系统为75个）大小为1920B的包在一个25Mbit/s的通道上传输。接下来的62个（1080i25系统为75个）大小

为1920B的包在第2个25Mbit/s的通道上传输。接下来的63个（1080i25系统为75个）大小为1920B的包在第3个25Mbit/s的通道上传输。最后62个（1080i25系统为75个）大小为1920B的包在第4个25Mbit/s的通道上传输。

188

16. 数字摄像机规范

1394商务协会（1394 Trade Association）为基于1394的数字视频摄像机编写了一个规范，其目的是为了避免完全实现IEC 61883规范的硬件和软件所需的高额费用。

定义了7种分辨率以及所支持的大量格式：

160 × 120	4：4：4 YCbCr
320 × 240	4：2：2 YCbCr
640 × 480	4：1：1、4：2：2 YCbCr，24位RGB
800 × 600	4：2：2 YCbCr，24位RGB
1024 × 768	4：2：2 YCbCr，24位RGB
1280 × 960	4：2：2 YCbCr，24位RGB
1600 × 1200	4：2：2 YCbCr，24位RGB

所支持的帧率有1.875帧/秒、3.75帧/秒、7.5帧/秒、15帧/秒、30帧/秒和60帧/秒。

同步包用于在1394网络上传输非压缩的数字视频数据。

参考文献

1. 1394-based Digital Camera Specification, Version 1.20, July 23, 1998.
2. Digital Transmission Content Protection Specification Volume 1(Information Version), July 25, 200.
3. Digital Visual Interface (DVI), April 2, 1999.
4. EBU Tech. 3267-E, 1992, *EBU Interfaces for 625-Line Digital Video Signals at the 4：2：2 level of CCIR Recommendation 601*, European Broadcasting Uion, June, 1991.
5. IEC 61883-1, 2003, *Consumer Audio/Video Equipment-Digital Interface-Part 1: General.*
6. IEC 61883-2, 1998, *Consumer Audio/Video Equipment-Digital Interface-Part 2: SD-DVCR Data Transmission.*
7. IEC 61883-3, 1998, *Consumer Audio/Video Equipment-Digital Interface-Part 3: HD-DVCR Data Transmission.*
8. IEC 61883-4, 1998, *Consumer Audio/Video Equipment-Digital Interface-Part 4: MPEG-2 TS Data Transmission.*
9. IEC 61883-5, 1998, *Consumer Audio/Video Equipment-Digital Interface-Part 4: SDL-DVCR Data Transmission.*
10. ITU-R BT.656-4, 1998, *Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4：2：2 level of Recommendation ITU-R BT.601.*
11. ITU-R BT.799-3, 1998, *Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4：4：4 level of Recommendation ITU-R BT.601 (Part A).*
12. ITU-R BT.1302, 1997, *Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4：2：2 level of Recommendation ITU-R BT.601.*
13. ITU-R BT.1303, 1997, *Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems Operating at the 4：4：4 level of Recommendation ITU-R BT.601 (Part B).*
14. ITU-R BT.1304, 1997, *Checksum for Error Detection and Status Information in Interfaces Conforming to ITU-R BT.656 and ITU-R BT.799.*
15. ITU-R BT.1305, 1997, *Digital Audio and Auxiliary Data as Ancillary Data Signals in Interfaces Conforming to ITU-R BT.656 and ITU-R BT.799.*
16. ITU-R BT.1362, 1998, *Interfaces For Digital Component Video Signals in 525-Line and 625-Line Progressive Scan Television Systems.*
17. ITU-R BT.1364, 1998, *Format of Ancillary Data Signals Carried in Digital Component Studio Interfaces.*
18. ITU-R BT.1365, 1998, *24-Bit Digital Audio Format as Ancillary Data Signals in HDTV Serial Interfaces.*
19. ITU-R BT.1366, 1998, *Transmission of Time Code and Control Code in the Ancillary Data Space of a Digital Television Stream According to ITU-R BT.656, ITU-R BT.799, and ITU-R BT.1120.*
20. ITU-R BT.1381-1, 2001, *Serial Digital Interface-based Transport Interface for Compressed Television Signals in Networked Television Production Based on Recommendations ITU-R BT.656 and ITU-R BT.1302.*

189

21. ITU-R BT.1577, 2002, *Serial Digital Interface-based Transport Interface for Compressed Television Signals in Networked Television Production Based on Recommendations ITU-R BT.1120.*
22. ITU-R BT.1616, 2003, *Data Stream Format for the Exchange of DV-based Audio, Data and Compressed Video over Interfaces Complying with Recommendation ITU-R BT.1381.*
23. ITU-R BT.1617, 2003, *Format for Transmission of DV Compressed Video, Audio and Data over Interfaces Complying with Recommendation ITU-R BT.1381.*
24. ITU-R BT.1618, 2003, *Data Structure for DV-based Audio, Data and Compressed Video at Data Rates of 25 and 50 Mbit/s.*
25. IRU-R BT.1619, 2003, *Vertical Ancillary Data Mapping for Serial Digital Interface.*
26. ITU-R BT.1620, 2003, *Data Structure for DV-based Audio, Data and Compressed Video at a Data Rate of 100Mbit/s.*
27. Kikuchi, Hidekazu, et al., *A 1-bit Serial Interface Chip Set for Full-Color XGA Pictures*, Society for information Display, 1999.
28. Kikuchi, Hidekazu, et al., *Gigabit Video Interface: A Fully Serialized Data Transmission System for Digital Moving Pictures*, International Conference on Consumer Electronics, 1998.
29. Open LVDS Display Interface (OpenLDI) Specification, v0.95, May 13,1999.
30. SMPTE 125M-1995, *Television-Component Video Signal 4 : 2 : 2—Bit-Parallel Digital Interface.*
31. SMPTE 240M-1999, *Television-Signal Parameters—1125-Line High-Definition Production Systems.*
32. SMPTE 244M-2003, *Television—System M/NTSC Composite Video Signals-Bit—Parallel Digital Interface.*
33. SMPTE 259M-1997, *Television—10-Bit 4 : 2 : 2 Component and 4F_{SC} Composite Digital Signals—Serial Digital Interface.*
34. SMPTE 260M-1999, *Television—1125/60 High-Difinition Production System—Digital Representation and Bit-Parallel Interface.*
35. SMPTE 266M-2002, *Television—4 : 2 : 2 Digital Component Systems—Digital Vertical Interval Time Code.*
36. SMPTE 267M-1995, *Television—Bit-Parallel Digital Interface-Component Video Signal 4 : 2 : 2 16 × 9 Aspect Ratio.*
37. SMPTE 272M-1994, *Television—Formating AES/EBU Audio and Auxiliary Data into Digital Video Ancillary Data Space.*
38. SMPTE 274M-2005, *Television—1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates.*
39. SMPTE 291M-1998, *Television—Ancillary Data Packet and Space Formatting.*
40. SMPTE 292M-1998, *Television—Bit-Serial Digital Interface for High-Difinition Television Systems.*
41. SMPTE 293M-2003, *Television—720 × 483 Active Line at 59.94 Hz Progressive Scan Production-Digital Representation.*

190

42. SMPTE 294M-2001, *Television—720 × 483 Active Line at 59.94 Hz Progressive Scan Production—Bit-Serial Interface.*
43. SMPTE 296M-2001, *Television—1280 × 720 Progressive Image Sample Structure, Analog and Digital Representation and Analog Interface—Bit-Serial Interface.*
44. SMPTE 305.2M-2000, *Television—Serial Data Transport Interface (SDTI).*
45. SMPTE 314M-1999, *Television—Data structure for DV-Based Audio, Data and Compressed Video—25 and 50Mb/s.*
46. SMPTE 326M-2000, *Television—SDTI Content Package Format (SDTI-CP).*
47. SMPTE 334M-2000, *Television—Vertical Ancillary Data Mapping for Bit-Serial Interface.*
48. SMPTE 334M-2000, *Television—540 Mbps Serial Digital Interface.*
49. SMPTE 348M-2000, *Television—High Data-Rate Serial Data Transport Interface (HD-SDTI).*
50. SMPTE 370M-2002, *Television—Data Structure for DV-Based Audio, Data and Compressed Video at 100 Mb/s 1080/60i,1080/50i,720p.*
51. SMPTE 372M-2002, *Television—Dual Link 292M Interface for 1920 × 1080 Picture Raster.*
52. SMPTE 396M-2003, *Television—Packet Format and Transmission Timing of DV-Based Data Streams over IEEE 1394.*
53. SMPTE RP-165-1994, *Television—Error Detection Checkwords and Status Flags for Use in Bit-Serial Digital Interfaces for Television.*
54. SMPTE RP-174-1993, *Bit-Parallel Digital Interface for 4 : 4 : 4 Component Video Signal (Single Link).*
55. SMPTE RP-175-1997, *Digital Interface for 4 : 4 : 4 Component Video Signal (Dual Link).*
56. SMPTE RP-168-2002, *Definition of Vertical Interval Switching Point for Synchronous Video Switching.*
57. SMPTE RP-188-1999, *Transmission of Time Code and Control Code in the Ancillary Data Space of a Digital Television Data Stream.*
58. SMPTE RP-208-2002, *Transport of VBI Packet Data in Ancillary Data packet.*
59. Teener, Michael D. Johns, *IEEE 1394-1995 High Performance Serial Bus*, 1394 Developer's Conference, 1997.
60. VESA DFP 1.0: Digital Flat Panel (DFP) Standard.
61. VESA Video Interface Port (VIP), Version 2, October 21, 1998.
62. VMI Specification, v1.4, January 30, 1996.

191

63. Wickelgren, Ingrid J., *The Facts about FireWire, IEEE Spectrum*, April 1997.

第7章　数字视频处理

除了需要对MPEG、NTSC/PAL和很多其他类型的视频进行编码和解码外，典型的视频系统通常还要进行很多其他的视频处理。

由于很多消费类显示器和大多数计算机显示器为逐行设备（非隔行的），因此隔行视频必须转换为逐行的（非隔行的）。逐行视频必须转换为隔行视频以驱动普通的模拟VCR或隔行TV，这就需要进行非隔行到隔行的转换。

很多计算机显示器支持每秒至少75帧的刷新速率。基于CRT的电视机刷新速率为每秒50或59.94（60/1.001）场（field）。对于平板电视而言，达到每秒120帧的刷新速率变得越来越普遍。对于基于电影（film-based）的压缩内容，视频源的帧率仅可以达到每秒24帧。这样，必须进行某种形式的帧率转换。

另外一个比较重要（not-so-subtle）的问题是视频缩放。SDTV和HDTV支持多种分辨率，然而显示器可能只有一种固定的分辨率。

alpha混合和色度抠像用于将多个视频信号或带有计算机生成文本和图形的视频混合起来。alpha混合确保在源视频之间的平滑交叉（smooth crossover），准许文本的亚像素定位，并限制源视频变换带宽以简化最终复合视频信号的编码。

由于没有完美的视频源，故即使是数字视频源，用户也总是希望可以对亮度、对比度、饱和度和色度进行控制。

192

7.1　舍入考虑

两个8位值相乘，生成的结果为16位。在某些情况下，必须将得到的乘积结果进行舍入处理得到较低精度的值（例如，16位舍入为8位、32位舍入为16位），这是为了满足高性价比的硬件实现。有多种舍入技术：截断、传统舍入、误差反馈舍入和动态舍入。

7.1.1　截断

截断就是在每个舍入操作过程中将任何小数部分的数据直接丢弃。其结果是，在只要很少的几个操作以后，就可能会引入严重的误差。在纯色（solid clor）区域这也可能会让人感觉到轮廓。

7.1.2　传统舍入

传统的舍入根据小数部分的数位来确定是舍（round down）还是入（round up）。如果小数部分为0.5或更大，那么就执行入操作——正数就应该变得更大，而负数则应该变得更小。如果小数部分小于0.5，那么就应该执行舍操作——正数应该变得更小，而负数则应该变得更大。

7.1.3　误差反馈舍入

误差反馈舍入服从规则“从来不丢弃任何东西”。误差反馈舍入的实现方式是，保存截断后的残差，并将其加入到下一个视频样本点中。这种方法用更不可见的类似噪声的量化误差代替了

简单截断所引起的轮廓效应。图7-1给出的是实现误差反馈舍入的例子。在这个例子中，16位的值通过误差反馈方法减少为8位的值。

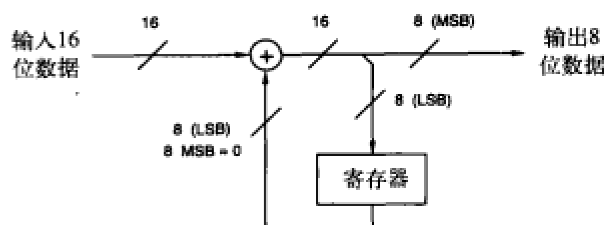


图7-1 误差反馈舍入

7.1.4 动态舍入

动态舍入这种方法（由Quantel公司注册的专利）根据丢弃小数位的权重对最低有效位（LSB）进行抖动（dither）处理。将初始的数据字分为两部分，一部分表示最终输出字的分辨率，一部分处理剩下的小数部分。小数数据与随机数生成器的输出进行比较，该随机数生成器在分辨率上与小数部分相等。比较器的输出为由小数部分的值进行加权的1位随机模式，并将作为加法器的进位。在任何情况下，输出字仅有一个最低有效位会改变，而且这种改变是随机的。图7-2所示的是实现动态舍入的例子。

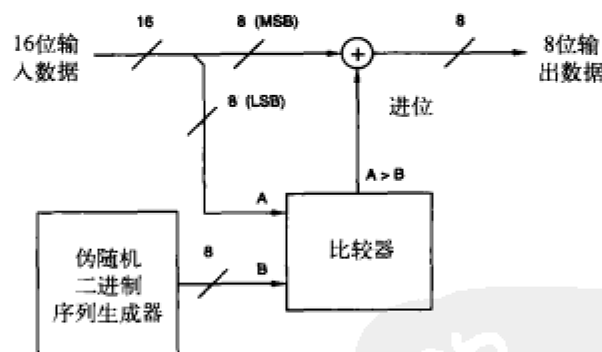


图7-2 动态舍入

7.2 SDTV-HDTV YCbCr转换

SDTV和HDTV应用具有不同的色度特征，如第3章所讨论。这样，当SDTV（HDTV）数据在HDTV（SDTV）显示器上显示时，就应该对YCbCr数据进行处理以弥补色度特征的差异。

7.2.1 SDTV到HDTV的转换

为了将Y_{601}CbCr（SDTV）数据转换到Y_{709}CbCr（HDTV），需要使用一个3×3的矩阵：

1	–0.11554975	–0.20793764
0	1.01863972	0.11461795
0	0.07504945	1.02532707

注意，在进行处理之前，必须去掉8位DC偏移（Y为16，CbCr为128），然后在处理完后再加在后面。

7.2.2 HDTV到SDTV的转换

为了将Y_{709}CbCr（HDTV）数据转换到Y_{601}CbCr（SDTV），需要使用一个3×3的矩阵：

1	0.09931166	0.19169955
0	0.98985381	−0.11065251
0	−0.07245296	0.98339782

注意，在进行处理之前，必须去掉8位DC偏移（Y为16，CbCr为128），然后在处理完后再加在后面。 194

7.3 4:4:4到4:2:2 YCbCr的变换

将4：4：4 YCbCr数据变换到4：2：2 YCbCr（图7-3）是数字视频中的一个通用功能。4：2：2 YCbCr是很多数字视频接口的基础，相对4：4：4来讲，实现4：2：2需要更少的连接。

饱和度逻辑应该包含在Y、Cb和Cr数据路径中，这样可以将8位数据的表示范围限制在1~254。值16和128用于在消隐期内生成合适的电平，如图7-3所示。

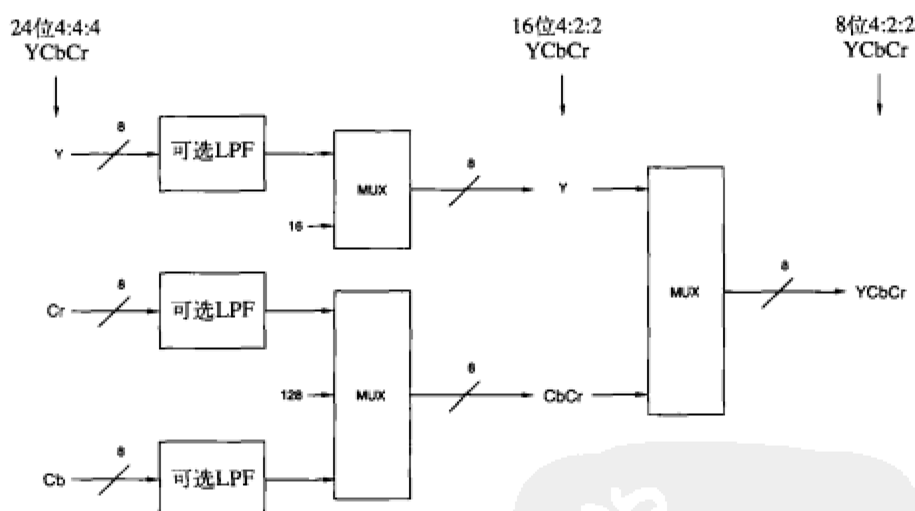


图7-3 4：4：4到4：2：2 YCbCr的变换

7.3.1 Y滤波

对Y进行滤波的低通滤波器（lowpass filter）的模板见图7-4和表7-1。

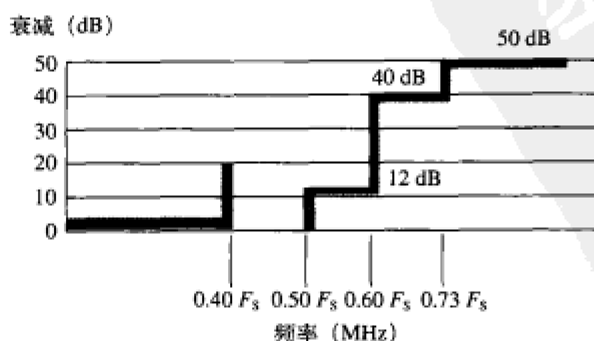


图7-4 Y滤波模板。F_S＝Y_1倍采样率

表7-1　Y滤波褶皱和群延迟限差。$F_S=Y_1$倍采样率。$T=1/F_S$

频率范围	典型的SDTV限差	典型的HDTV限差
	带通褶皱限差	
0~0.40 F_S	由±0.01dB增加到±0.05dB	±0.05dB
	带通群延迟限差	
0~0.27F_S	由0增加到±1.35ns	±0.075T
0.27 F_S~0.40 F_S	由±1.35ns增加到±2ns	±0.110T

因为滤波过程可能有很多级联（cascaded）的变换（可以想象最多可达10级），所以设计的滤波器要保持非常小的误差容限，以避免视觉伪影的积累。因为在通过多级（mulitple stage）处理后，滤波会放大实际信号与平滑幅度和群延迟响应的偏离程度。例如，如果采用的滤波器在1MHz时引起的噪声为−1dB，而在1.3MHz时为−3dB，那么在经过了4级变换后（假设每级2个滤波器），总的响应在1MHz处为−8dB，而在1.3MHz处为−24dB。

尽管陡峭截止会导致Y边沿处的振铃效应（ringing），但如果群延迟性能足够好，视觉伪影应该是最小的。当级联多个滤波器进行操作时，通带平坦性和群延迟特性就是非常重要的。与陡峭截止相耦合的带通容限使得模板匹配非常困难（从某种程度上讲是不可能的）。因此，通常有一个倾向就是放松通带准确度，但最好的方法就是降低截止速度，并保持通带尽可能平坦。

195

7.3.2　CbCr滤波

Cb和Cr经过了低通滤波和抽取。在标准的设计中，低通滤波器和抽取滤波器可以组合成一个单一的滤波器，并且单个滤波器可以通过复用的方式同时应用于Cb和Cr。

与Y滤波的情况一样，Cb和Cr低通滤波要求具有陡峭的截止频率，这样可以防止在反复变换的过程中导致分辨率损失的累积。然而，对于Cb和Cr而言，由于截止频率较低，陡峭截止频率产生的振铃效应要比Y的情况明显得多。

图7-5和表7-2给出的是Cb和Cr滤波器的一个模板。

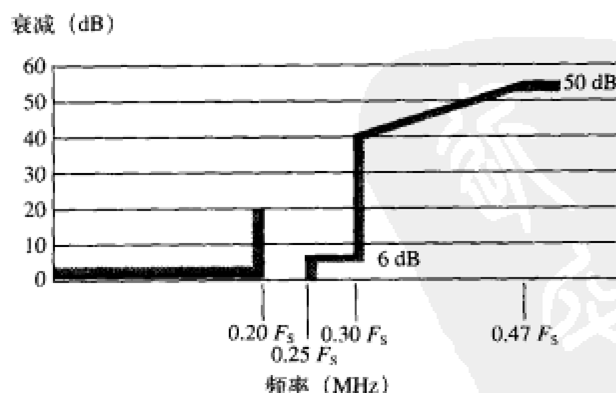


图7-5　将采样速率从4∶4∶4变换为4∶2∶2的数字滤波器的Cb和Cr滤波器模板。$F_S=Y_1$倍的采样速率

表7-2　CbCr滤波褶皱和群延迟限差。$F_S=Y_1$倍的采样速率。$T=1/F_S$

频率范围	典型的DSTV限差	典型的HDTV限差
	通带褶皱限差	
0~0.20 F_S	0dB增加到±0.05dB	±0.05dB
	通带群延迟限差	
0~0.20 F_S	通过设计，使得延迟失真为0	

由于色差信号的混叠没有那么明显，因此，在二分之一采样频率处的衰减仅为6dB。使用斜对称（skew-symmetric）的频率响应以二分之一采样频率通过−6dB点有一个优点：这使得数字滤波器的交替系数为0，几乎使抽头数减半，并且对于Cb和Cr信号也可以使用一个单一的滤波器。使用横向滤波器的优点是能够提供完美的线性相位响应，消除了对群延迟进行校正的必要。

与Y滤波器的情况一样，通带的平坦性和群延迟特性是非常重要的，最好的方法仍然是降低截止速度，尽可能保持通带平坦。

7.4 显示增强

7.4.1 亮度、对比度、饱和度（颜色）和色调（色彩）

在YCbCr颜色空间进行处理简化了亮度、对比度、饱和度和色调控制的实现，如图7-6所示。图中同时还展示了复用器，使用复用器可以输出黑屏、蓝屏和彩条。

这样的设计应当确保不会出现上溢或下溢反转误差，能够有效地将饱和颜色值限定在0~255。

1. Y处理

将数据Y减去16就可以将黑色电平定位于0处。这消除了DC偏移，因而调整对比度就不会改变黑色电平。由于输出数据Y的值可以小于16，这时就应该支持Y值为负的情况。

对比度（或图像或白色电平）控制是通过将YCbCr数据乘以一个常数来实现的。如果没有调整Cb和Cr，那么任何时候，彩色漂移都可能导致对比度发生变化。典型的8位对比度调整范围是0~1.992×。

亮度（或黑色电平）控制是通过将数据Y加上或减去一个数来实现的。亮度控制是在对比度控制之后实现的，这是为了避免由于调整对比度而引入不断变化的DC偏移值。典型的8位亮度调整范围是−12~+127。

最后，要将数据Y加上16以便将黑色电平定位于16的位置。

2. CbCr处理

从Cb和Cr上减去128就可以将它们的取值范围定位于0附近。

色调（或色彩）控制是通过混合Cb和Cr数据来实现的：

$$Cb' = Cb\cos\theta + Cr\sin\theta$$

$$Cr' = Cr\cos\theta - Cb\sin\theta$$

其中θ是期望的色调角。典型的8位色度调整范围是−30°~+30°。

196~198

饱和度（或颜色）控制是通过将Cb和Cr都乘一个常数来实现的。典型的8位饱和度调整范围是0~1.992×。在图7-6所给的例子中，对比度和饱和度值乘在一起是为了减小CbCr数据路径上的乘法器数量。

最后，在Cb和Cr上都加上128。

很多显示器也分别对红色、绿色、蓝色、青色、黄色和洋红色使用独立的色度和饱和度控制。这使得在生成图像时可以对图像进行调整，以更好地匹配显示器的特性。

7.4.2 彩色瞬时改善

YCbCr过渡应该是对齐的。然而，Cb和Cr过渡通常是减缓的，时间上是有偏差的，这是由于色差信息较窄的带宽导致的。

通过对一致的Y过渡进行监视，可以合成Cb和Cr更快的水平和垂直过渡。也可以将小的预发射（pre-shoot）和后发射（after-shoot）加到Cb和Cr信号上。

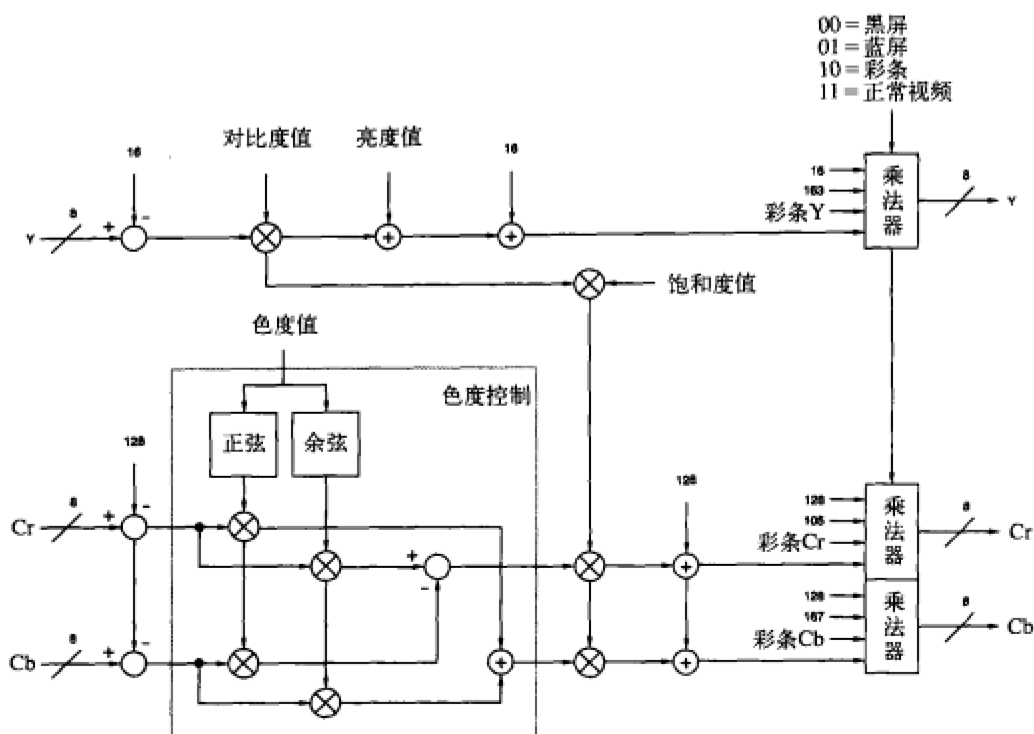


图7-6　色调、饱和度、对比度和亮度控制

这样，新的Cb和Cr边沿就会与Y边沿对齐，如图7-7所示。

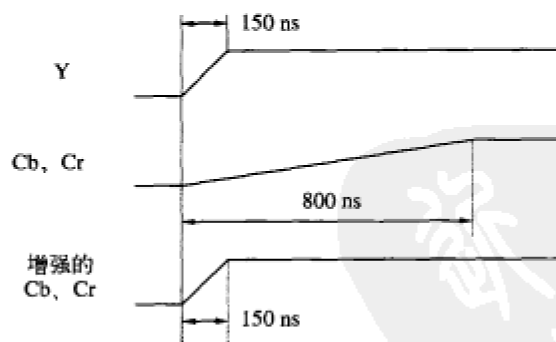


图7-7　颜色瞬时改善

通常，显示器利用这种技术来提供看起来对比更加显著的画面。

7.4.3　亮度瞬时改善

在这种情况下，Y水平和垂直过渡缩短了，也可以将小的预发射（pre-shoot）和后发射（after-shoot）加到Y上，以手工方式对图像进行锐化。

显示器通常使用这种技术来提供看起来对比更加明显的画面。

7.4.4　锐化

画面锐度的显著增加可以通过增加高频亮度信息的幅度来实现。

如图7-8所示，可以使用具有选择性增益的简单带通滤波器（也可以称为峰值滤波器）。出现最大增益的频率通常可以选择在颜色副载波频率处或大约2.6MHz处。典型情况下，在滤波器之后使用除噪电路（coring circuit）来降低电平噪声。

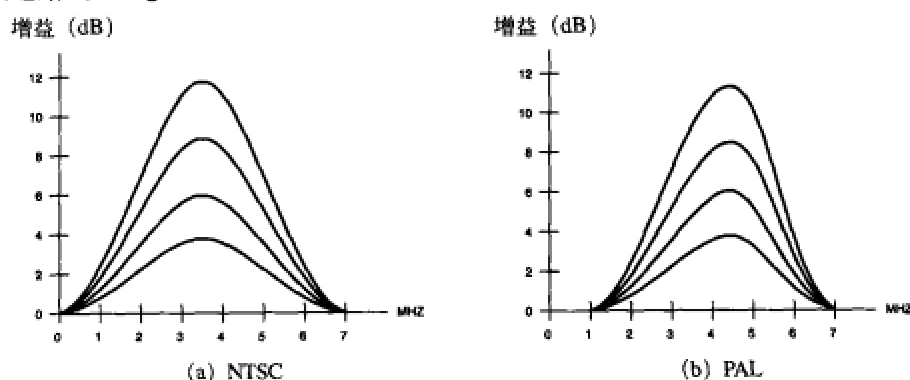
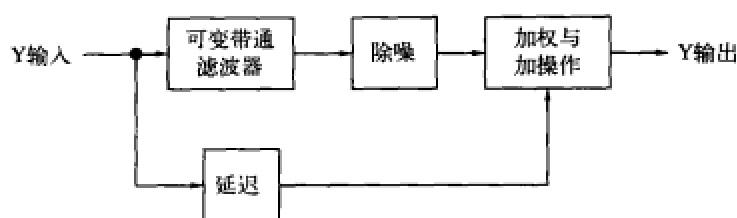
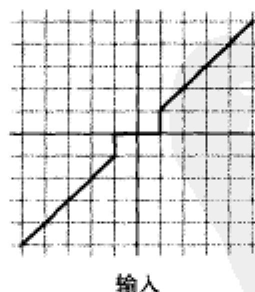


图7-8 简单的可调锐度控制

图7-9给出的是一个更复杂的锐化控制电路。使用具有可调增益的可变带通滤波器增加高频亮度。除噪功能（典型值为±1 LSB）消除了低电平噪声。然后将修改的亮度加到初始的亮度信号上。



(a) 典型实现



(b) 除噪功能

图7-9 更复杂的锐度控制

除了可选增益外，也应该支持可选的高频衰减。很多电视通过提升高频增益来改善图像的显示锐度。尽管可以关闭对电视的锐度控制，但是这会影响模拟广播的图像质量。

199
200

7.4.5 蓝色延展

蓝色延展（blue strench）增加了白色和白色附近颜色的蓝色值，这是为了使得白色看起来更亮。当采用蓝色延展时，应该只能处理指定范围内的颜色。

具有最大值的约80%或更多的Y值的颜色的饱和度值低，落在CbCr平面内的白色检测区域内，

其蓝色分量增加约4%（蓝色增益因子），而红色分量减少同样的数量。对于更加复杂的设计，白色检测区域和蓝色增益因子可能依赖于颜色的Y值和饱和度值。

过渡边界可以用在白色检测区域附近，这样随着颜色偏离白色检测区域边界，就可以逐渐地减小蓝色增益因子。这可能会妨碍蓝色延展区域与没有进行颜色延展的区域之间的硬过渡。如果颜色落在过渡边界区域内，那么它就使用蓝色增益因子的一小部分对蓝色进行延展，并且随着离检测区域边界的距离的增加，这部分占蓝色增益因子的比例会不断降低。

7.4.6　绿色增强

当绿色电平值较低时，绿色增强创建了一个更丰富、更饱和的绿色。显示器通常使用这种技术来提供看起来更绿的草、树木等。当采用绿色增强时，应该处理仅位于特定颜色范围内的颜色。

具有低绿色饱和度值并且落在CbCr平面内绿色检测区域内颜色的饱和度增加了。有些设计不是将绿色检测区域的中心定位于绿色轴（图9-28中的241°）附近，而是使用约213°作为绿色检测轴，因此，同样的设计也可以很容易地用于实现肤色校正（skin tone correction）。

简单的实现就可以让最大饱和增益（约1.2倍）出现在绿色检测轴上，随着与绿色检测轴距离值的增加，饱和度增益减小到1倍。对于更加复杂的设计，绿色检测区域和最大饱和度增益可能依赖于颜色的Y值和饱和度的值。

有些显示器也使用这种技术来实现蓝色增强，用它来使得天空看起来更蓝。

7.4.7　动态对比度

使用动态对比度（也称为自适应对比度增强），根据图像的内容，图像暗的部分和亮的部分之间的差异通过手工方式得到增强。显示器通常使用这种技术来改善它们的对比度比率。

主要以暗为主的图像中的亮色通过使它们变得更亮（白色延展）而得到增强。典型情况下，这是通过使用直方图信息来修正gamma曲线的上面部分而实现的。

主要以较亮画面为主的图像中的暗色通过使这些图像更暗（黑色延展）而得到增强。典型情况下，这是通过使用直方图信息来修正gamma曲线的下面部分实现的。

对于中等亮度的图像，两种技术都可以采用。

较小的gamma校正调整也可以用于暗色和亮色之间的颜色，这可以让图像具有更多细节和对比更显著。

201~202

7.4.8　颜色校正

在源视频与显示器使用的视频之间，RGB色度通常会存在细微的差异。这会导致红色、绿色和蓝色不是完全地准确。

颜色校正可以用在源视频上，对显示器的特征进行补偿，使得显示器可以显示出更加准确的红、绿和蓝色。

颜色校正的一个替代方式是进行颜色扩展，这可以利用现代显示器所具有的更加强大的颜色再生能力。颜色扩展可以使绿色更绿，蓝色更蓝，等等。实现颜色扩展的一个常见技术是对于每个主色、辅色和肤色都使用独立的色度和饱和度控制。

7.4.9　色温校正

在没有校准过的电视机中，色温（白色）的变化是基于亮度值的。

D_{65}（由大多数视频标准定义的白色点）的色温是6500°K。6500°K以上的色温更蓝（冷）；6500°K以下的色温更红（暖）。

很多电视在工厂装配时具有非常高的平均色温（7000~8000°K），这是为了强调电视机的亮度。观众可以选择2到3个出厂设置（暖、冷等）或观看模式（电影、运动等），它们可以作为色温的参考。“冷”设置更亮（就像你在正午的阳光下看到的一样），更适合白天的场景（比如运动赛事），这是由于增强了亮度。“暖”设置更柔和（就像你在室内环境下柔和的灯光下看到的一样），这对于观看电影或在黑暗环境下观看比较合适。

色温可以通过一个3×3的矩阵乘法器处理YCbCr或R′G′B′数据。10个寄存器（从10到100 IRE，每10 IRE步使用1个）为3×3矩阵乘法器提供9个系数。寄存器中的值由校准过程决定。作为中间IRE电平的YCbCr或R′G′B′值可以通过插值的方式来确定。 203

7.5 视频混合与图形覆盖

混合视频信号可能像在两个视频源之间切换一样简单。如果混合后的视频是用于在计算机监视器上显示，那么这已经足够了。

对于大多数的应用，应该使用alpha混合（alpha mixing）技术。alpha混合技术也可以用于过渡到某个特定的颜色（如黑色）或从某个特定的颜色过渡到别的图像，还可以用于将计算机生成的文本和图形覆盖到视频信号上。

如果要将视频编码为复合视频，那么必须使用alpha混合。否则，在视频源切换点处（如在计算机生成的文本和图形的边缘附近）可能会出现振铃效应和模糊。这是由于在NTSC/PAL编码器中，颜色信息经过了低通滤波的缘故。如果滤波器具有陡峭的截止频率，那么快速的颜色过渡将会产生振铃效应。另外，在视频处理的路径上的某处，亮度信息的带宽可能被限制在4~5MHz的范围内，这将减缓亮度信息的过渡。

在数学上，对alpha值进行了归一化处理，其取值范围为0~1，alpha混合可以按如下方式实现：

$$out = (alpha_0)(in_0) + (alpha_1)(in_1) + \cdots$$

在这种实现中，每个视频源都有它自己的alpha信息。alpha信息之和可以不等于1（单位增益）。

图7-10表示的是两个YCbCr视频信号的混合，每个视频信号都有自己的alpha信息。由于YCbCr使用的是偏移二进制表示，偏移（Y为16，Cb和Cr为128）在视频信号混合之前就被消除掉了。在完成混合之后，再将偏移量加到混合好的信号上。注意，也可以直接对两个4：2：2 YCbCr视频流进行处理，没有必要将它们变换到4：4：4 YCbCr格式进行混合，然后再将结果变换回4：2：2 YCbCr格式。

当只有两个视频源参与混合并且alpha_0＋alpha_1＝1（实现交叉渐变器）时，在数学上，单一的alpha值可以按照下面的方式使用：

$$out = (alpha)(in_0) + (1 - alpha)(in_1)$$

当alpha＝0时，输出结果就等于in_1视频信号；当alpha＝1时，输出结果就等于in_0视频信号。当alpha值在0到1之间时，两个视频信号就按照比例与alpha值相乘，然后将结果相加。

将前面的公式进行扩展和调整，就可以看出两通道混合器是如何利用单个乘法器实现的：

$$out = (alpha)(in_0 - in_1) + in_1$$

过渡到特定的颜色或从特定的颜色过渡到其他的画面，都可以通过将一个输入源设置为固定的颜色来实现。

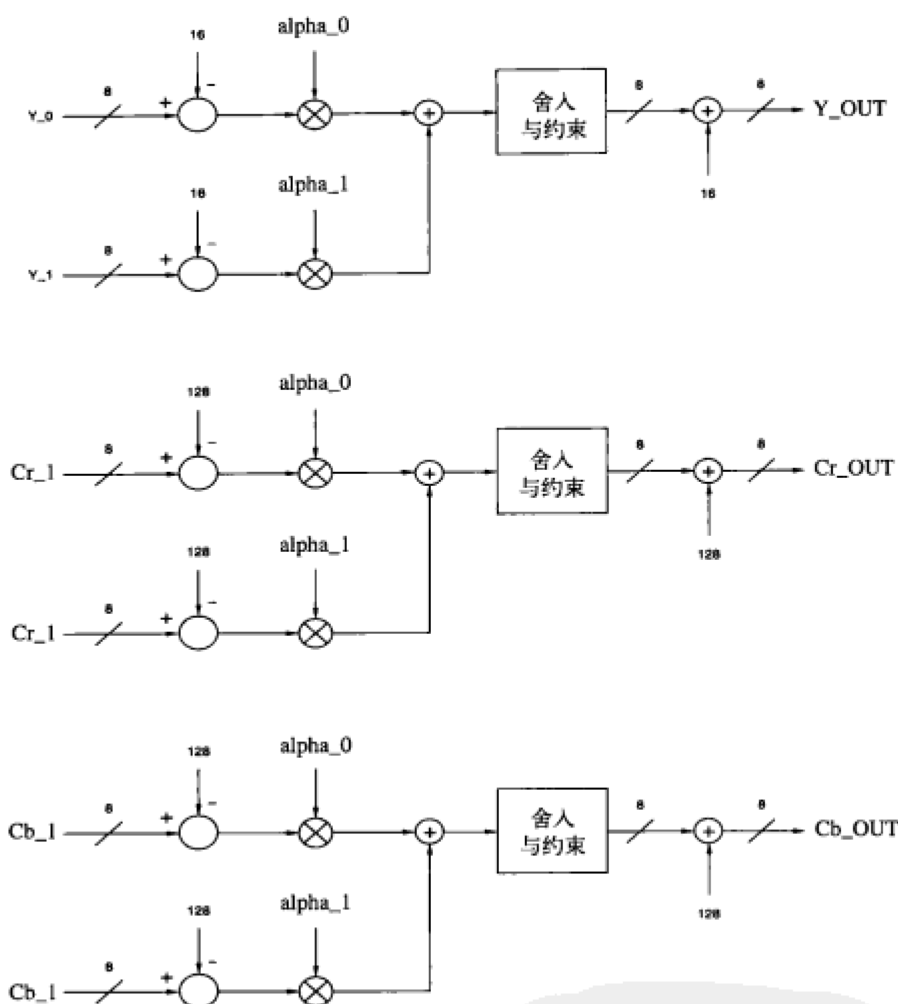


图7-10　混合两个YCbCr视频信号，每个都有自身的alpha通道

图7-11表示的是使用单个alpha通道将两个YCbCr源混合在一起。图7-12和图7-13表示的是混合两个R'G'B'视频源（R'G'B'的取值范围是0~255）。图7-14和图7-15表示的是混合两个数字复合视频信号。

使用alpha值的计算机图形系统有一个共同的问题，就是帧缓存可能包含预处理过的R'G'B'或YCbCr数据，也就是说，在帧缓存中的R'G'B'或YCbCr数据已经与alpha相乘过了。假定alpha（A）的值为0.5，未处理过的R'G'B'A值，对于白色为（255, 255, 255, 128）；预处理过的R'G'B'A值，对于白色为（128, 128, 128, 128）。所以，任何从帧缓存中接受的R'G'B'或YCbCr数据都应该能够处理两种格式。

204~205

通过调整alpha值，由慢到快的交叉过渡是可以实现的，图7-16所示样本点之间alpha值相差太大就会导致快速的交叉过渡；相差较小就会导致较慢的交叉过渡。如果将alpha混合用于特效（如擦除效果），那么切换点（其中每个视频源使用50%）的精度必须能够调整到小于一个样本点，这样才能够确保画面的平滑过渡。通过控制alpha值，可以有效地将切换点定位于任何位置，如图7-16a所示。

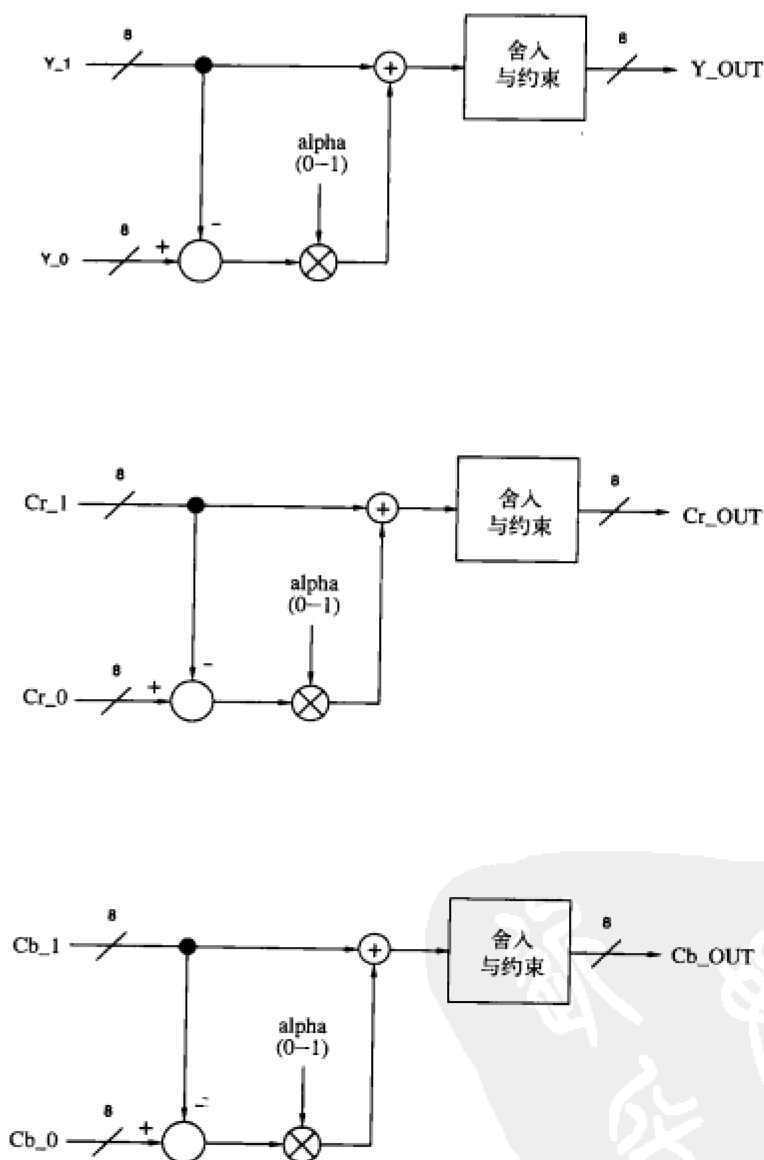
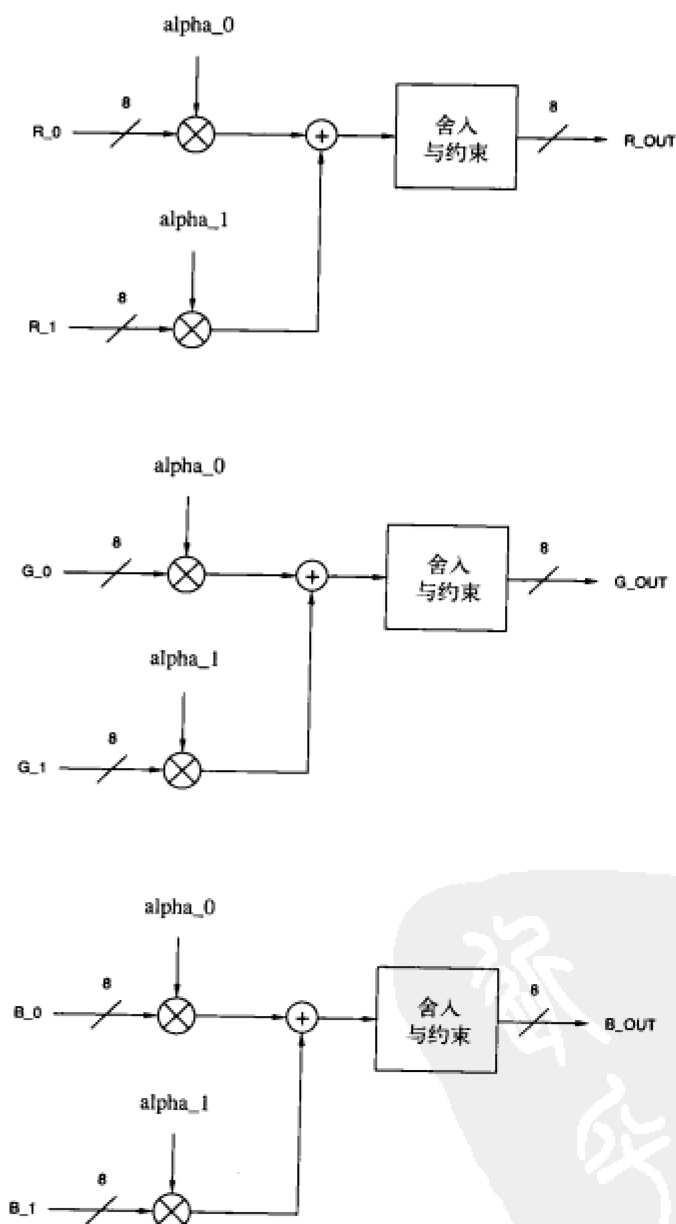


图7-11 两个YCbCr视频信号混合（交叉过渡）的简化。使用单一的alpha通道

206



207

图7-12 混合两个RGB视频信号（RGB的取值范围是0~255），每个视频信号都有自己的alpha通道

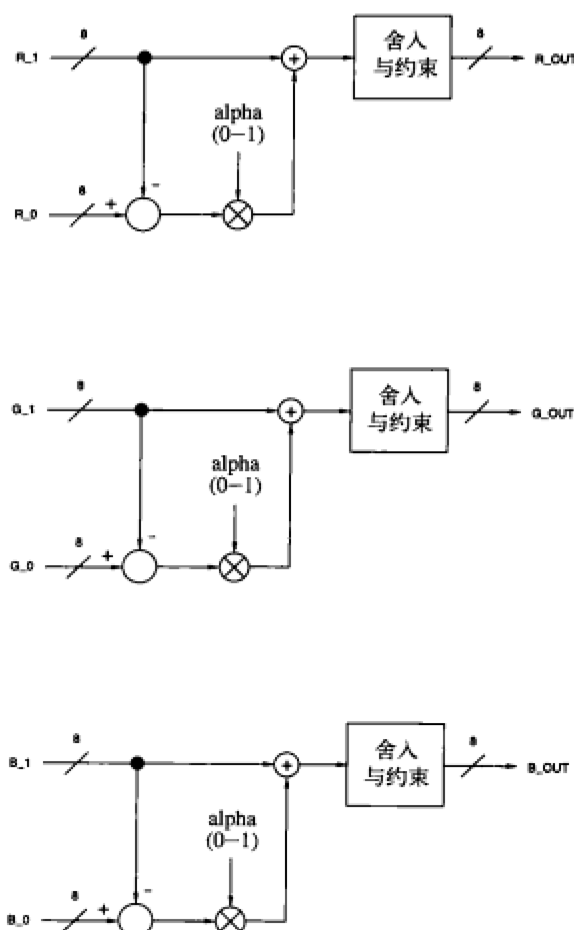


图7-13 两个RGB视频信号（RGB的取值范围是0~255）混合（交叉过渡）的简化，使用单一的alpha通道 208

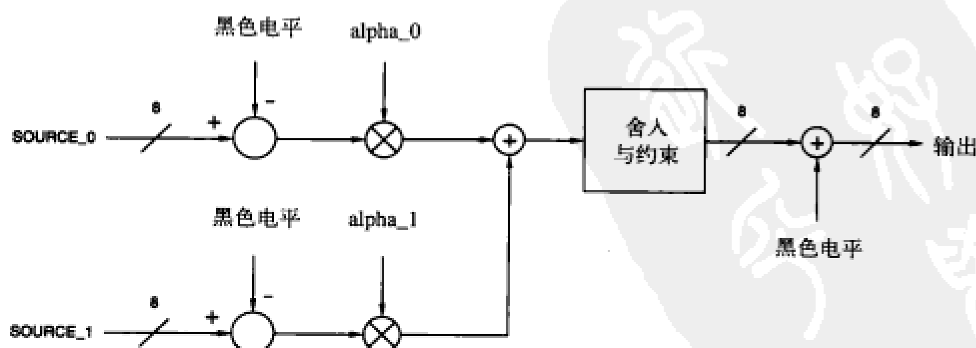
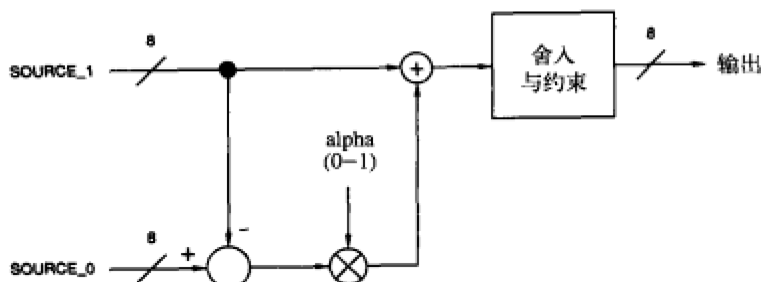
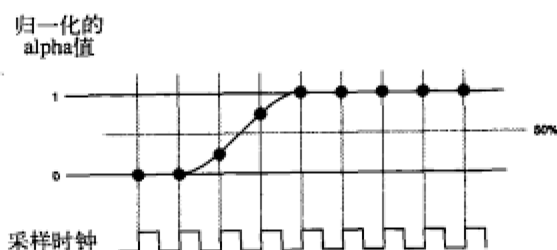


图7-14 混合两个数字复合视频信号，每个视频信号都有自己的alpha通道

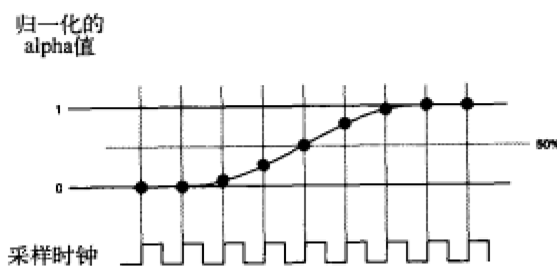


209

图7-15 两个数字复合视频信号混合（交叉过渡）的简化。每个视频信号都有自己的alpha通道



(a)



(b)

图7-16 控制alpha值以实现（图a）快或（图b）慢抠像。在图a中，有效的切换点位于两个样本之间。在图b中，过渡时间范围更宽，并且是与采样时刻对齐的

210

让字符生成器控制alpha的输出就可以将文本覆盖到视频上。将其中一个输入源设置为固定的颜色，文本就会呈现出这种颜色。

注意，对于在处理前从Y通道减去16（黑色电平）的那些设计，在减操作执行后应该支持负的Y值。这样才可以使设计通过现实世界的检验，并对具有最小伪影的视频信号进行测试。

7.6 亮度和色度抠像

抠像需要指定一个期望的前景色，包含前景色的区域就用背景图像替换。另外，可以指定任意大小或形状的任何区域，在这个区域内（或外）的前景区域都用背景图像替换。

7.6.1 亮度抠像

亮度抠像涉及指定期望的前景亮度值，包含高于（或低于）抠像值的亮度值的前景区域用背景图像替换。

另外，这种硬抠像的实现可以用软抠像代替，这是通过给前景图像指定两个亮度值来实现的：Y_H和Y_L（$Y_L<Y_H$）。对于将背景抠像到白色的前景区域，高于Y_H的前景亮度值（Y_{FG}）用背景图像替换；而低于Y_L的Y_{FG}值包含前景图像。对于Y_H和Y_L之间的值，在前景和背景图像之间执行线性混合。这种操作可以表示为：

如果$Y_{FG}>Y_H$，$K=1$=仅取背景图像；

如果$Y_{FG}<Y_L$，$K=0$=仅取前景图像；

如果$Y_H \geqslant Y_{FG} \geqslant Y_L$，$K=(Y_{FG}-Y_L)/(Y_H-Y_L)$=前景与背景的混合。

由1中减去K，可以生成抠像到黑色前景区域的新亮度抠像信号。

图7-17表示的是两个YCbCr视频源的亮度抠像。尽管典型情况下，色度抠像使用压缩技术来消除前景图像中的信息，但当作为Cb和Cr幅度的亮度抠像通常与亮度值没有关系时，并不这样做。

图7-17 两个YCbCr视频信号的亮度抠像

图7-18表示的是两个R′G′B′视频源的亮度抠像，这更适用于计算机图形。Y_{FG}可以通过下面的公式获得：

$$Y_{FG}=0.299R'+0.587G'+0.114B'$$

在这种应用中，红色和蓝色数据被忽略了，这会导致Y_{FG}仅等于绿色数据。

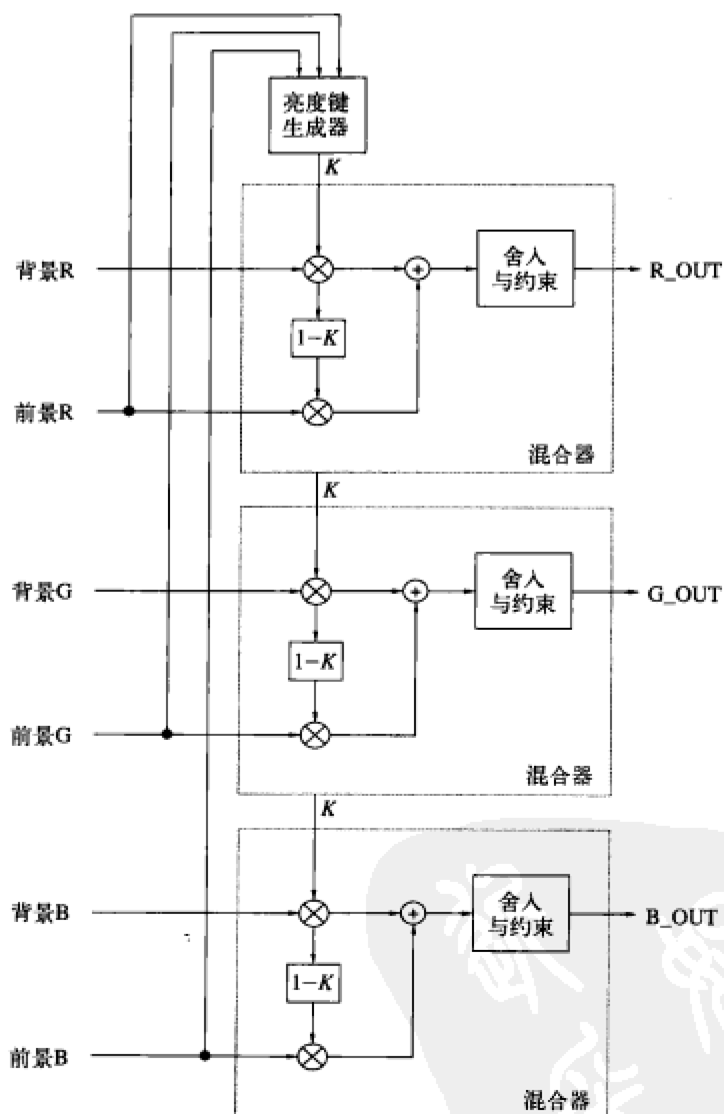


图7-18　两个RGB视频信号的亮度抠像。RGB的取值范围是0~255

211
~
213

图7-19表示的是两个数字复合视频源之间实现亮度抠像的技术。

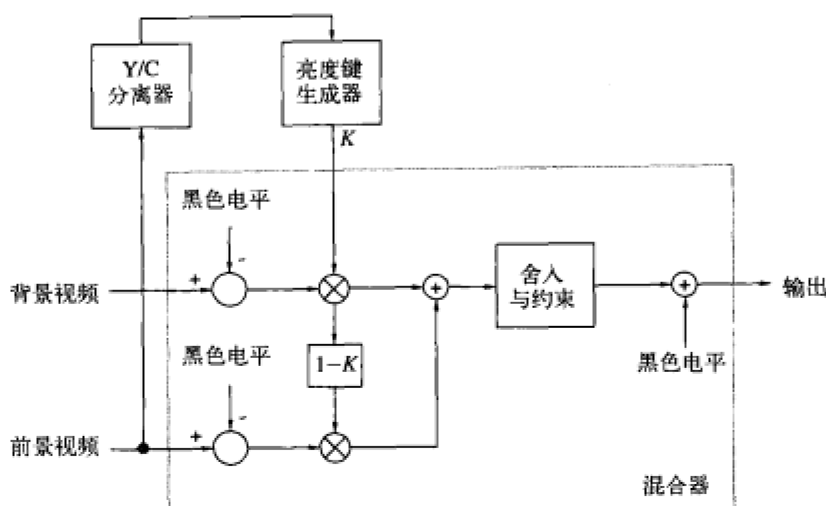


图7-19 两个数字复合视频信号的亮度抠像

7.6.2 色度抠像

色度抠像需要指定期望的前景抠像颜色；包含抠像颜色的前景区域用背景图像替换。Cb和Cr用于指定键颜色；亮度信息可以用于增加色度抠像功能的真实性。尽管分量混合减少了伪影，但两个视频源的实际混合可以在分量或复合视频域内完成。

早期的色度抠像电路简单地在前景和背景图像之间进行硬或软切换。除了对前景图像中保持的精细细节的数量进行限制外，通过透明或半透明前景对象，背景是不可见的，并且前景阴影不会出现在包含背景图像的区域内。

建立线性抠像器可以将前景和背景图像按照键值确定的比例组合在一起，这导致在包含背景图像的区域内前景图像会被衰减。尽管准许前景图像对象出现透明，但是对前景内保持的细节精细度还是有限制的。如果不进行额外的处理，那么前景形成的阴影就不会出现在包含背景图像的区域内，背景图像特定区域的亮度值必须减小，以产生前景对象投影形成的阴影效果。 [214]

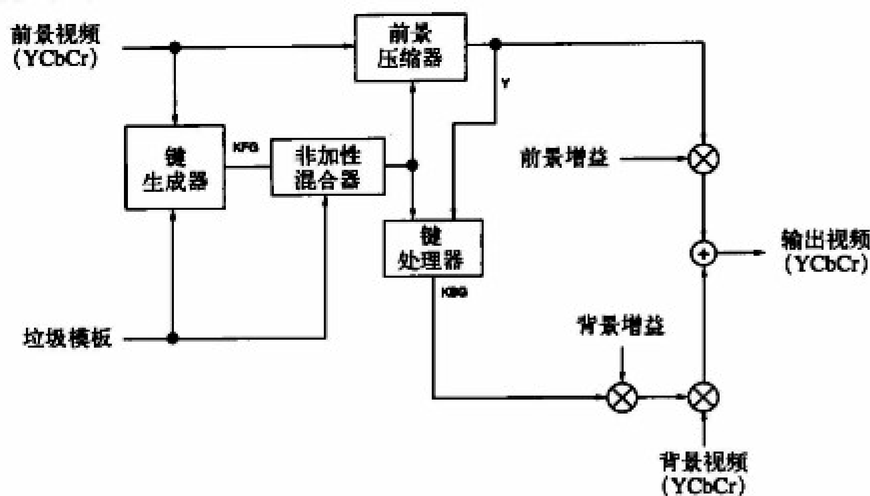
如果除了前景对象投影形成的阴影外，前景场景使用的蓝色或绿色衬底亮度均匀，那么背景上的效果将是前景对象投影形成的阴影，这个被称为阴影色度抠像或亮度调制的过程使得可以对亮度值进行调节，这种调节是与前景场景中蓝色或绿色衬底的亮度值是成比例的。因此，保持谱的高光部分，可以使透明或半透明前景对象具有更加真实的抠像。

注意，绿色背景现在使用更多，因为它可以降低色度噪声。

色度抠像器的能力也限于处理不需要切换到背景图像的靠近键颜色的前景颜色，另外一个问题可能是前景对象是淡蓝色（bluish tint），这是从蓝色衬底上反射的蓝光或摄像机镜头里发生的漫反射造成的。色度溢出（chroma spill）是很难消除的，因为溢出的颜色不是起初的键颜色；还会进行一些混合，这将稍微改变初始的键颜色。

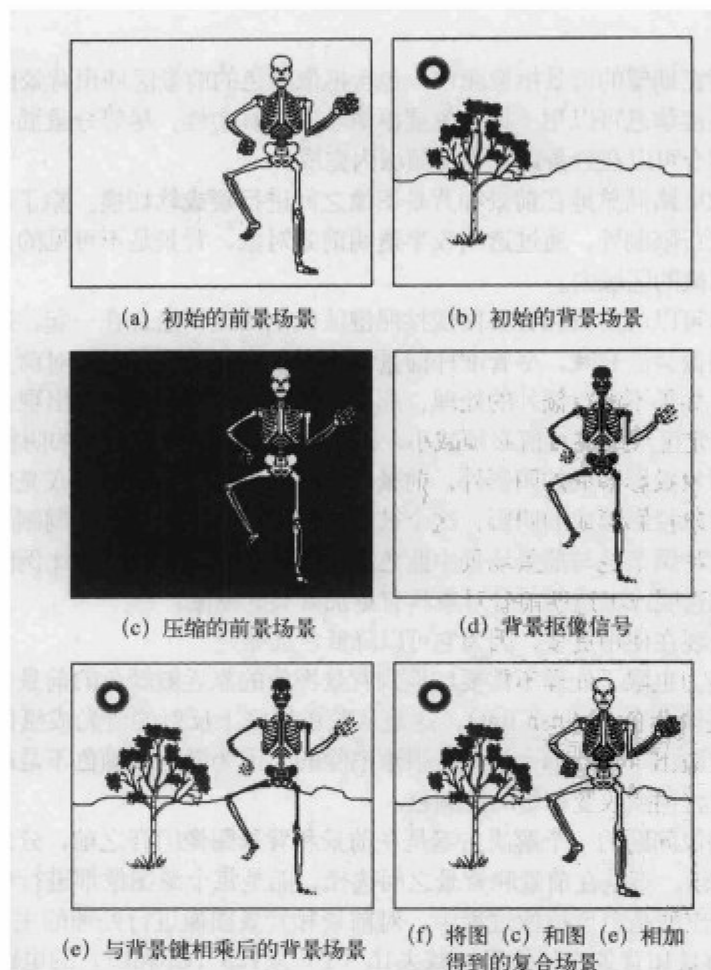
对于很多色度抠像问题的一个解决方案是在前景和背景图像组合之前，分别单独地对它们进行处理，如图7-20所示。不是在前景和背景之间选择，而是每个源图像都进行单独处理，然后再进行组合。图7-21给出的是色度抠像过程中，对前景和背景图像进行处理的主要步骤。在图7-20中没有给出起初从前景和背景视频信号中减去16（Y）或128（Cb和Cr）的电路，也没有在最后的输出加法器之后加16（Y）和128（Cb和Cr）的电路。任何没有消除的DC偏移都将被前景和背

景增益因子所放大或削弱，并使黑色电平发生漂移。



215

图7-20 典型的分量色度键电路



216

图7-21 色度抠像过程中的主要处理步骤

前景键（K_{FG}）和背景键（K_{BG}）信号的取值范围为0~1。垃圾模板键信号（术语模板来自于电影制片企业）强制混合器以两种方法之一输出前景源。

第一种方法是按照K_{FG}的增加成比例地减小K_{BG}。这种方法有一个优点：它最小化了前景与背景交接处的黑边。

第二种方法是强制背景中模板键的所有非0值为黑色，并将前景插入背景中形成的洞中。这要求有一个清除函数来消除黑色电平周围的噪声，因为直接相加的过程会使这种噪声影响背景图像。

使用非加性混合器将垃圾模板加到前景键信号（K_{FG}）上。非加性混合器（NAM）取两幅图像中较亮的样本点，样本点逐个地产生键信号。对于任何产生自身抠像信号的视频源来讲，模板是理想的，如字符生成器等。

键生成器监视前景的Cb和Cr数据，生成前景抠像信号K_{FG}。选择一个期望的键颜色，如图7-22所示。对前景中的Cb和Cr数据进行归一化（生成了Cb′和Cr′），并旋转一个θ角，这样就生成了X和Z数据，使得正X轴尽可能近地穿过期望的键颜色。典型情况下，θ可以按照增1的方式发生变化，最优的色度抠像出现在X轴通过键颜色时。

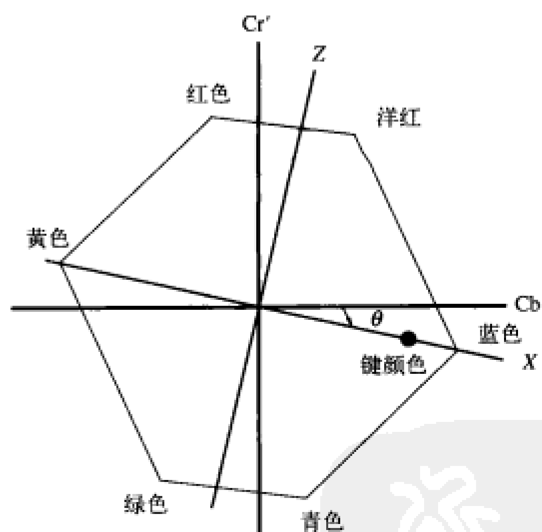


图7-22 将归一化的Cb和Cr（Cb′和Cr′）轴旋转角θ来获得X和Z轴，使得X轴通过期望的键颜色（在本例中为蓝色）

X和Z是按照下面的公式由Cb和Cr推导而来的：

$$X = Cb'\cos\theta + Cr'\sin\theta$$

$$Z = Cr'\cos\theta - Cb'\sin\theta$$

由于对Cb′和Cr′进行了归一化，其取值范围在±1之间，X和Z的取值范围就是0~1。

前景抠像信号（K_{FG}）是由X和Z来生成的，其取值范围是0~1：

$$K_{FG} = X - \{|Z|/[\tan(\alpha/2)]\}$$

$$K_{FG} = 0，前提是X < \{|Z|/[\tan(\alpha/2)]\}$$

其中的α为接受角（acceptance angle），关于正X轴中心对称，如图7-23所示。在接受角以外，K_{FG}总是设为0。在接受角以内，K_{FG}的幅度随前景颜色趋近于键颜色和饱和度的增加而线性增加。在接受角以内的颜色由前景处理器进一步进行处理。

前景压缩器通过实现$X=X-K_{FG}$来减少前景颜色信息，并且键颜色被压缩到黑色电平上。为了在$K_{FG}=0$时避免对Cb和Cr进行处理，前景压缩器需要执行下面的操作：

$$Cb_{FG}=Cb-K_{FG}\cos\theta$$

$$Cr_{FG}=Cr-K_{FG}\sin\theta$$

其中Cb_{FG}和Cr_{FG}是经过键颜色压缩后的前景Cb和Cr值。早期的实现通过将Cb和Cr与Cb_{FG}信号的剪切版本相乘来压缩前景信息。但是，由于乘法和剪切过程，这会产生带内混叠分量，并且会在键颜色边界处产生不自然的边界。

217

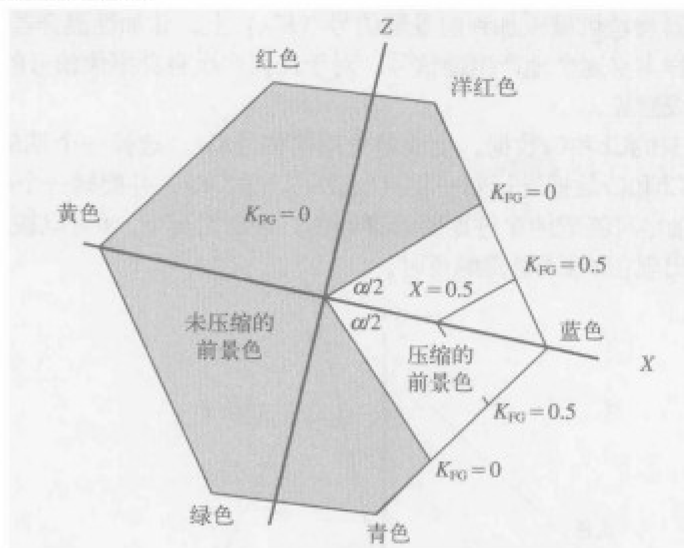


图7-23　前景键值和接受角

如果不进行额外的处理，那么只要Cb_{FG}和Cr_{FG}是准确地定位于X轴上的，这两个分量就设为0。因为噪声或光照引起的色度变化将会导致前景区域没有被完全地压缩，所以，压缩角的设置要关于正X轴中心对称。典型情况下，从最小的0度角到最大的三分之一接受角（α）的范围内，压缩角（β）是可以配置的。任何落在这个压缩角范围内的CbCr分量都要设置为0。图7-24给出了压缩角的使用方式。

前景亮度在归一化处理之后的取值范围为0~1，它可以通过下面的方式进行压缩：

$$Y_{FG}=Y'-y_S K_{FG}$$

$$Y_{FG}=0，前提是y_S K_{FG}>Y'$$

这里的y_S是一个可编程改变的值，可以用它来调整K_{FG}，使得在键颜色区域内的黑色电平处可以对K_{FG}进行剪切。

前景压缩器也消除了由于色度溢出引起的所需前景区域上的键颜色边缘，这是通过去掉所需前景对象的污点来实现的。

Ultimatte®通过测量蓝色与绿色之间的差异来改善这个过程，由于蓝色衬底总不是纯蓝色的，在前景对象中可能存在高电平的蓝色。纯蓝在自然界是很少见的，大多数自然界中的蓝色具有比红色更多的成分。由于这个原因，可以对红色、绿色和蓝色电平进行监测，以区分所需前景对象中的蓝色衬底和蓝色。

如果蓝色和绿色之间的差异足够大，那么所有三种颜色都设置为0，就可以产生黑色，这就是在包含蓝色衬底的区域内需要进行的处理。

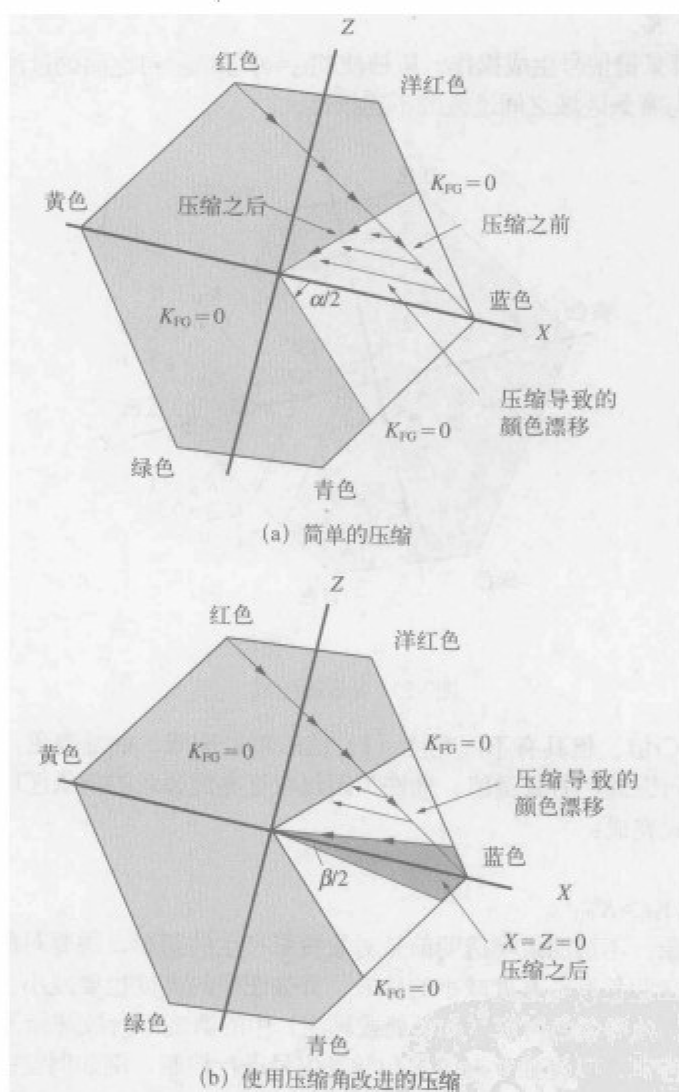


图7-24 从红色前景对象到蓝色键颜色渐变过程的压缩角操作

如果蓝色和绿色之间的差异不大，只要绿色没有超过红色，就将蓝色设为绿色电平。这种技术可以消除由于蓝色衬底引起的淡蓝色，从而能够在前景中再生出自然的蓝色。例如，白色前景通常是由相同电平的红色、绿色和蓝色组成的。如果白色区域受到键颜色（在此为蓝色）的影响，那么将会产生淡蓝色——蓝色电平将大于红色或绿色电平。因为绿色没有超过红色，所以使蓝色电平等于绿色电平，消除了淡蓝色。

然而，这需要付出一定的代价。前景色中的洋红色会变为红色。可以使用绿色衬底，但是在这种情况下，前景中的黄色就会被修改。通常，可以逐步放松钳位（clamping）来增加洋红色区域中的蓝色成分。

键处理器生成初始的背景键信号（K'_{BG}），在背景可见的地方，用它来消除背景图像的区域。在期望的背景图像区域中，将K'_{BG}调整为0，而在没有衰减的背景区域中，将K'_{BG}调整为单位数据。由背景键信号（K_{BG}）可以生成K'_{BG}，具体过程是使K_{BG}通过偏移（K_I）和增益（K_G）调整，然后在0和单位值内进行截取，就可以得到K'_{BG}：

$K'_{BG} = (K_{FG} - K_I)\ K_G$

图7-25表示的是背景键信号生成操作。应该使$K'_{BG}=0$与$K'_{BG}=1$之间的过渡尽可能宽，这样可以尽可能最小化前景与背景区域之间过渡的不连续性。

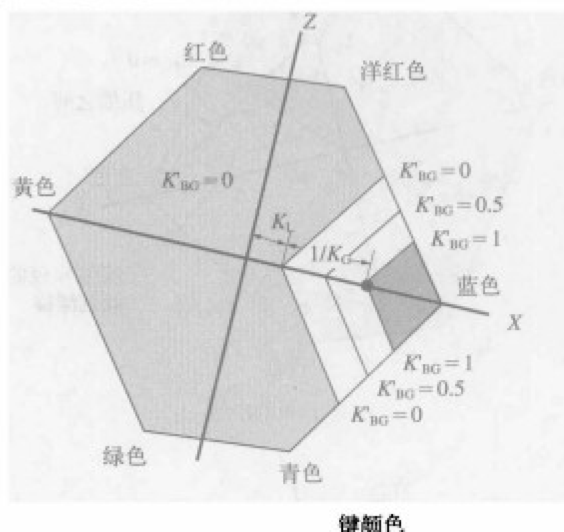


图7-25 背景键生成

218~219

对于包含同样CbCr值、但具有不同亮度（Y）值的前景区域，如键颜色，随着前景亮度电平的增加，键处理器也可以减小背景键值，准许关闭包含更亮键颜色的前景区域中的背景，如亮蓝色。可以用下面的公式完成：

$K_{BG} = K'_{BG} - y_c Y_{FG}$

$K_{BG}=0$，如果 $y_c Y_{FG} > K'_{BG}$

为了处理前景对象、不透明或半透明前景对象投影产生的阴影，需要对前景图像蓝色衬底的亮度电平进行监测。在蓝色衬底亮度减小的地方，背景图像的亮度也要减小。必须对背景图像亮度的减小量进行控制，使得蓝色衬底（如裂缝或脚注）中的瑕疵不会被解释为前景中的阴影。

可以使用额外的控制，以便独立对前景和背景信号进行控制，例如调整前景的对比度使得它匹配背景，或以各种方式隐退前景（如过渡到背景，使前景对象消失，或隐退到黑色以生成轮廓）。

在计算机中，可能有相对较慢和平滑的边缘——尤其是涉及平滑隐退的边缘。因为平滑边缘在色度抠像过程中容易失真，所以，在这种环境下，通常采用宽的抠像过程。在宽抠像过程中，抠像信号开始于图形对象边缘之前。

220~221

复合色度抠像

在有些情况下，分量信号（如YCbCr）无法直接得到。对于这种情况，可能要使用复合色度抠像，如图7-26所示。

为了检测色度键颜色，必须对前景视频源进行解码，这样可以减小Cb和Cr颜色微分信号。然后用抠像信号K_{FG}对两个复合视频源进行混合。垃圾模板键信号通过减小K_{FG}强制混合器输出背景源。

由于存在颜色带宽不足的问题，使用复合视频信号的色度抠像经常导致不真实的抠像。因此，复合色度抠像处理过的图像缺少精确的细节，并且在边界上可能会出现光晕。

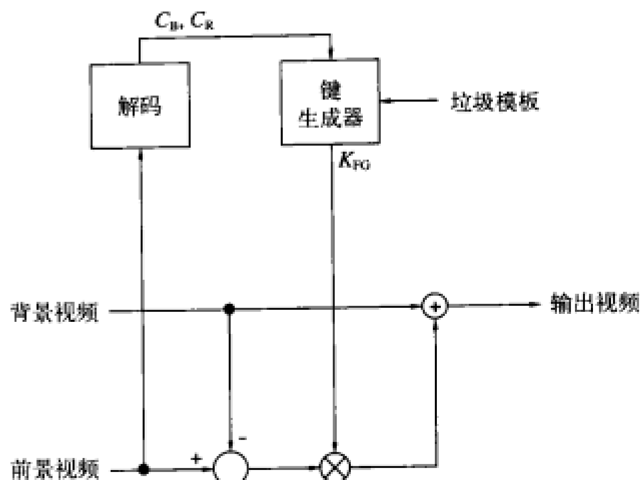


图7-26 典型的复合色度键电路

7.6.3 超黑色和亮度抠像

视频系统也可以利用超黑色（superblack）或亮度抠像（luma keying）。在低于消隐电平（模拟视频）或黑色电平（数字视频）的这个特定范围内取值的前景视频区域，用背景视频信息将其替换。

222

7.7 视频缩放

对于所有的视频分辨率（表7-3），在几乎每个解决方案中，缩放（scaling）通常都是必要的。

表7-3 消费类显示器和广播视频源的公共有效分辨率

显示器		SDTV视频源		HDTV视频源
704 × 480	640 × 480	704 × 360[1]	704 × 432[1]	1280 × 720
854 × 480	800 × 600	480 × 480	480 × 576	1440 × 816[2]
704 × 576	1024 × 768	528 × 480		1440 × 1040[3]
854 × 576	1280 × 768	544 × 480	544 × 576	1280 × 1080
1280 × 720	1366 × 768	640 × 480		1440 × 1080
1280 × 768	1024 × 1024	704 × 480	704 × 576	1920 × 1080
1920 × 1080	1280 × 1024		768 × 576	

1. 4：3显示器上16：9的Letterbox模式。
2. 1920 × 1080、16：9显示器上2. 35：1的变形模式。
3. 1920 × 1080、16：9显示器上1.85：1的变形模式。

当生成在SDTV上显示的对象时，计算机用户必须关注文本大小、线条的粗细等事情。例如，在1280 × 1024计算机显示器上可读的文本可能在SDTV显示器上是不可读的，这是由于涉及大量的缩小操作。当将细的水平线转换为隔行SDTV时，该水平线可能要么完全消失，要么以25或29.97Hz的频率闪烁。

注意，缩放必须在分量视频上进行（如R'G'B'和YCbCr）。复合颜色视频信号不能直接进行缩放，这是因为它提供了彩色副载波相位信息，在缩放后将变得没有意义。

通常，在输出采样点之间的空间可以通过目标递增（tarinc）值来定义：

tarinc = I/O

其中I和O是输入（I）和输出（O）采样点的数目，可以是水平方向上的，也可以是垂直方向上的。

通过将公式调整为下面的形式，可以使第一个和最后一个输出采样点与第一个和最后一个输入采样点对齐：

223

tarinc = $(I-1)/(O-1)$

7.7.1 像素丢弃与复制

像素丢弃与复制（pixel dropping and duplication）也称为“最近邻（nearest neighbor）”缩放，这是因为只使用了与输出采样点最近的输入采样点。

缩小最简单的形式是像素丢弃，其中在水平和垂直方向上，每n个样本点中就有m个被丢弃。典型情况下，使用Bresenham画线算法（在计算机图形学的书中有描述）的一个修改版来确定丢弃哪些采样点。

简单的放大操作可以通过像素复制来完成，其中在水平和垂直方向上，每n个像素中就有m个像素被复制。与缩小操作一样，用Bresenham画线算法的修改版确定复制哪些采样点。

在实际中，是不推荐使用像素丢弃或复制的缩放操作的，因为这会导致视觉伪影，并会引入混叠成分。

7.7.2 线性插值

使用线性插值可以改善缩放图像视频质量。当输出采样点落在两个输出样本点之间时（水平或垂直方向上），输出样本点是在两个输入样本点之间进行线性插值计算出来的。但是，将图像缩小到原图像的一半以下时，仍然会导致某些样本点被删除。

图7-27表示的是将16：9的图像在垂直方向上进行缩放，以适应在4：3的显示器上播放。通常使用的是一个简单的双线性滤波器，如图7-28a所示。对两个源样本点L_n和L_n+1加权，然后将它们加在一起形成一个目的样本点D_m。

$$D_0 = 0.75L_0 + 0.25L_1$$
$$D_1 = 0.5L_1 + 0.5L_2$$
$$D_2 = 0.25L_2 + 0.75L_3$$

然而，如图7-28a所示的一样，这将导致不均匀的线间隔，而不均匀的线间隔可能导致视觉上的伪影。图7-28b表示的是垂直滤波，这使得输出的线在空间分布上更加均匀：

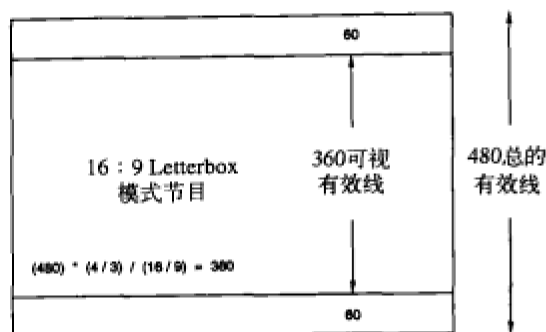
$$D_0 = L_0$$
$$D_1 = (2/3)L_1 + (1/3)L_2$$
$$D_2 = (1/3)L_2 + (2/3)L_3$$

线性插值器是一个较差的带限滤波器。它不必要地消除掉了过多的高频细节，在奈奎斯特极限之上保留了太多的能量，这都将导致混叠效应。

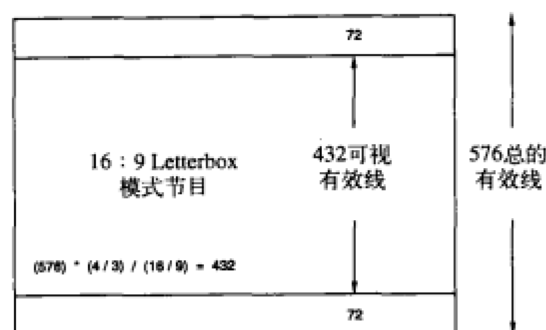
7.7.3 抗混叠重采样

最好的方法是保证在水平和垂直方向上，频率成分与图像大小成比例地缩放。

图7-29给出的是抗混叠重采样过程的基本原理。由A对输入数据进行向上采样，然后进行低通滤波以去掉插值过程产生的图像频率。滤波器B对信号进行带限滤波，以去掉在重采样过程B中导致混叠的频率成分。B/A的值决定了缩放因子。

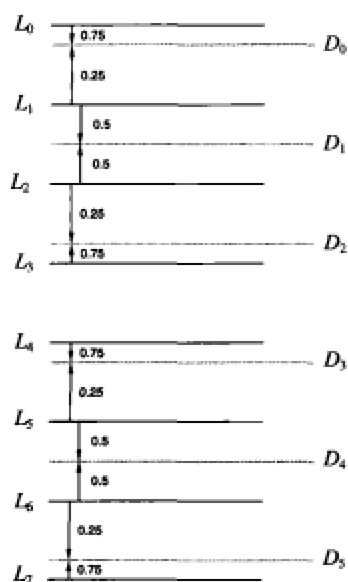


(a) 480线的系统

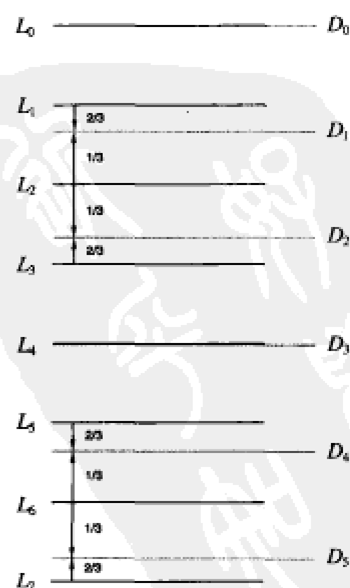


(b) 576线的系统

图7-27 对16：9的图像垂直进行缩放以适应在4：3的显示器上播放



(a) 结果为不均匀的空间分布



(b) 结果为均匀的空间分布

图7-28 将16：9的图像进行75%的缩放以适应在4：3的显示器上播放

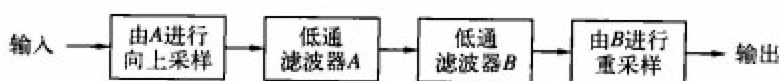


图7-29　抗混叠重采样器的一般结构

通常将滤波器A和B组合成一个单一的滤波器。滤波器的响应基本确定了插值的质量。理想的低通滤波器具有非常平坦的通带，在最低采样频率（输出或输入）的二分之一处有陡峭的截止，在阻带上具有非常高的衰减。但是，由于这样的滤波器会在陡峭的边界处产生振铃效应，因此通常期望复制出通带的顶部。这可以制作出稍微柔和但是具有更不显著的振铃效应的图像。

224～226

滤波器的通带褶皱（ripple）和阻带衰减可以测量缩放质量，但是振铃效应的主观影响意味着平坦的通带可能没有想象的那么好。大量的阻带衰减的存在几乎总是一件好事。

通用的重采样结构有3个基本形式。每个形式都以各种方式结合了图7-29中的组件。

一种方法是采用可变带宽抗混叠滤波器，在其后为组合插值器/重采样器。这种情况下，对于每个缩放因子，滤波器需要新的系数。因为缩放因子发生了变化，图像质量也可能发生改变。另外，如果使用线性插值，那么总体响应就比较差。但是，滤波器系数是时变的，并且不存在增益问题。

第二种方法是采用组合滤波器/插值器，紧接其后为重采样器。通常，插值的阶数n越高，总的响应也越好。滤波器变换函数的中心总是与新输出样本的顶端对齐。利用每个缩放因子，对滤波器变换函数进行延展或压缩，使得它保持与n个输出样本的顶端对齐。这样，滤波器的系数和所使用的输入样本的个数就会随着每个新输入样本和缩放因子而变化。需要动态地对增益进行归一化，这是为了保证滤波器系数的和总等于1。

第三种方法是采用插值器，在其后为组合滤波器/重采样器。插入样本0对输入数据进行插值，使得插值后的数据为输入和输出速率的公倍数。使用低通有限脉冲响应（FIR）滤波器对插值后的数据进行滤波，以便将样本插入到以0填充的缝隙中，然后在所要求的位置进行重采样。通常使用“多阶段（polyphase）”滤波器可以实现这种设计，随着输入和输出样本相对位置的变化，这种滤波器会切换其系数。

7.7.4　显示缩放示例

图7-30到图7-38分别给出了在4：3和16：9的显示器上显示16：9和4：3的图像的各种缩放的例子。



图7-30　16：9视频源的例子



图7-31　对16：9的内容进行缩放以适用4：3显示器：“正常”或平移和扫描（pan-and-scan）模式。导致某些16：9的图像内容被忽略（灰色显示的区域）



图7-32 对16：9的内容进行缩放以适用4：3显示器：Letterbox模式。全部的16：9节目都可见，在显示器的顶部和底部都出现了黑色条



图7-33 对16：9的内容进行缩放以适用4：3显示器："挤压（squeezed）"模式。在水平方向上将整个的16：9节目进行压缩以适应在4：3的显示器上显示，这样得到的将是一幅变形的图像



图7-34 4：3视频源的例子



图7-35 对4：3的内容进行缩放以适用16：9显示器："正常"模式。16：9显示器的左边和右边部分没有使用，因此显示为黑色或灰色



图7-36 对4：3的内容进行缩放以适用16：9显示器："宽屏"模式。整个图像在水平方向上进行了线性缩放以适应整个16：9的显示器，如果没有采用变形的视频内容，那么在屏幕上显示的图像将发生变形



图7-37 对4：3的内容进行缩放以适用16：9显示器："放大（zoom）"模式。4：3图像的顶部和底部被删除，然后进行缩放以适应整个16：9的显示屏

如何显示图像是由用户偏好和图像宽高比同时决定的。例如，当在4：3的显示器上播放16：9的视频时，很多用户喜欢让裁减过的图像填充整个显示屏（如图7-31所示），而不喜欢采用会产生黑色或灰色条纹的Letterbox模式（如图7-32所示）。

另外，有些显示器错误地假设到达这些显示器YPbPr输入上的任何逐行视频信号都来自于

"变形的"视频源。其结果是，当不应该采用缩放的时候，它们却在水平方向上将逐行的16：9节目放大了25%。所以，对于机顶盒，包含"16：9（压缩的）"模式是有益的，这种模式在水平方向上将逐行16：9的节目缩小了25%，这样就可以补偿16：9的显示器在水平方向上进行的放大处理。



图7-38 对4：3的内容进行缩放以适用16：9显示器："全景（panorama）"模式。在水平方向上，对图像的左右25%的边进行非线性缩放，以适应16：9的显示器，在图像的左右边缘发生了变形

7.8 扫描速率变换

227
~
231

在很多情况下，需要进行某些形式的扫描速率变换（也称为时间速率变换、帧率变换或场率变换）。多标准的模拟VCR和扫描变换器使用扫描速率变换来实现各种视频标准之间的变换。计算机通常以大约75Hz的非隔行方式运行，然而却需要显示50Hz和60Hz的隔行视频。而数字电视机可以同时支持多种帧率。

注意，处理必须在分量视频信号上进行（如R'G'B'或YCbCr）。复合彩色视频信号不能直接进行处理，这是由于它提供了颜色副载波相位信息，对复合视频信号进行处理以后，颜色副载波相位信息就会失去意义。

7.8.1 帧或场丢弃和复制

简单的扫描速率变换可以通过每*N*个场（field）丢弃或复制一个场来实现。例如，将60Hz的扫描速率变换为50Hz隔行视频的操作，利用一个场存储，可以通过每6个场丢弃一个场来实现，如图7-39所示。

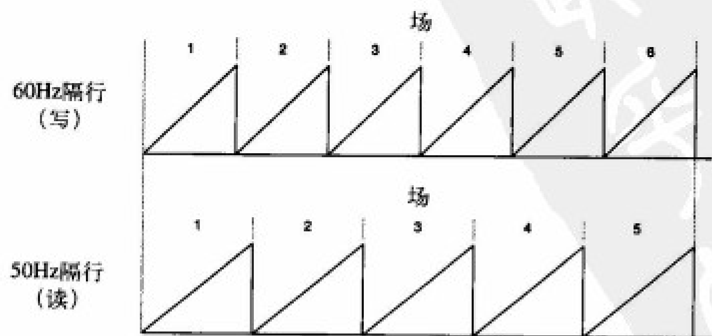


图7-39 使用单个场存储，通过每6个场就丢弃一个来实现60Hz到50Hz的变换

这种技术的不足在于观众可以看到画面发生抖动移动，或运动抖动（motion judder）。另外，有些视频解码产品仅使用最上面的场来实现从60Hz到50Hz的转换，这就损失了垂直方向上的分

辨率。

当进行非整数扫描速率变换时，会出现最糟糕的伪影。例如，当某些帧显示3次时，而其他的帧只显示2次。在这种情况下，观众将观察到重叠或模糊对象。因为人脑以连续的帧跟踪对象，期望看到一个正常的位置序列，当对走走停停的对象运动进行协调时就会出现麻烦。其结果是，观众会错误地得出结论，认为有两个对象在同时运动。

232↓233

7.8.2 时域插值

当需要生成希望的帧率时，这种技术就会根据初始帧来生成新帧。应该同时使用已经输入的和即将输入的帧的信息，这样就可以最优地处理对象的出现与消失。

图7-40给出了使用时域插值（temporal interpolation）实现从50Hz到60Hz的变换过程。对于50Hz视频的每5个场，在60Hz视频中则有6个场。

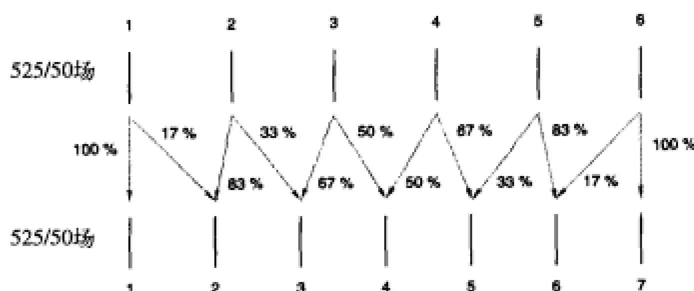


图7-40 没有进行运动补偿的情况下，使用时域插值将帧率由50Hz转换为60Hz

在两个源对齐以后，将两个50Hz的相邻场混合在一起来生成新的60Hz场。这种技术用在某些便宜的标准变换器中实现50Hz与60Hz标准之间的变换。注意，这里没有使用运动分析。因此，如果帧率为50Hz的摄像机在水平方向上平移经过一个狭长的垂直对象，那么每6个60Hz的场，就会看到这个对象一次，并且对于中间的5个场，将会看到两个对象，其中一个淡入，而另一个则淡出。

1. 50Hz到60Hz转换示例

图7-41给出了扫描速率变换器，该变换器实现了垂直方向上的插值，然后进行时域插值。图7-42给出了图7-41中设计的谱表示（spectral representation）。

现在的很多设计将垂直插值和时域插值结合起来形成一个单一的设计，如图7-43所示，对应的谱表示见图7-44。这个例子先使用了垂直插值，然后进行时域插值。如果是先进行时域插值，然后再进行垂直插值，那么所需的场存储大小将减半。但是，线存储的数目将由4增加到8。

在上面的每个例子中，第一个插值过程必须产生一个中间、高分辨率的逐行格式以避免隔行分量，因为隔行分量会干扰第二个插值过程。不论在垂直方向还是水平方向上，仅使用两个场的线混合的插值方法是不够的，因为插值过程不能补偿隔行线的时间偏移。

2. 运动补偿

使用时域插值的、具有更高质量的扫描速率变换器包含了运动补偿，这样就可以使伪影最小化。其结果是运动显得非常平滑和自然，图像显得更清晰（sharper），不会出现运动抖动（motion judder）。

针对扫描速率变换的运动估计不同于MPEG所使用的运动估计。在MPEG中，目标是通过搜索连续帧中区域之间高度的相关性来最小化移位帧差异。得到的运动向量不一定对应于真实的运

动向量。

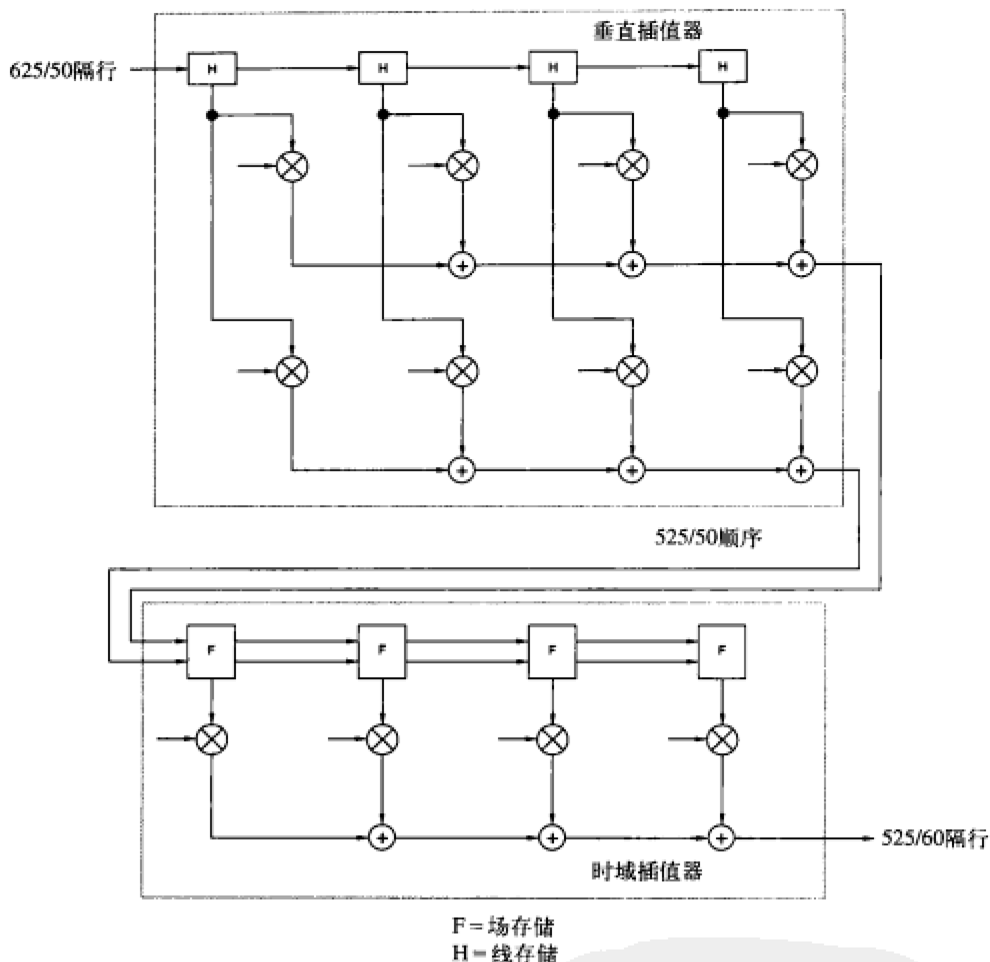


图7-41 典型的50Hz到60Hz帧率变换，使用垂直插值，紧接其后进行时域插值

对于扫描速率变换，确定真实的运动信息以实现正确的时域插值很重要。插值应该能够承受不正确的运动向量以避免引入讨厌的伪影，因为伪影是我们要采用技术尽可能消除的。有几种原因可能导致不正确运动向量，如跟踪运动的时间不足、运动向量超出范围和混叠引起的估计困难等。

3. 100Hz隔行电视的例子

234~235

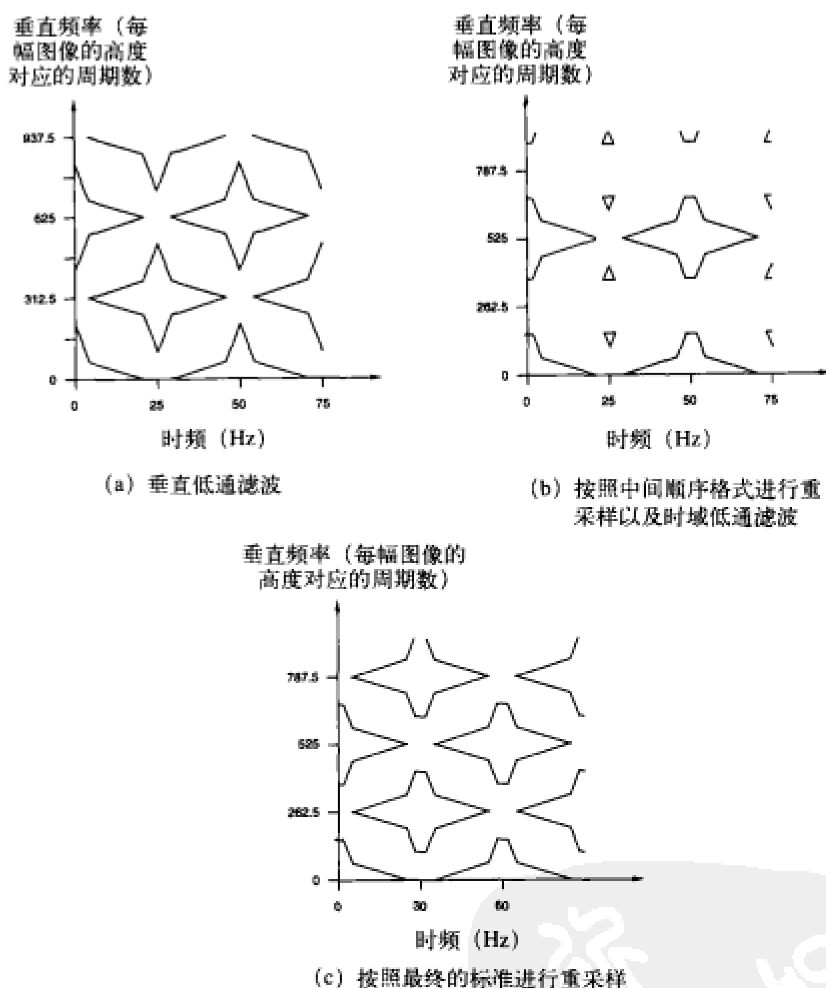
标准的50Hz隔行电视每秒钟显示50场。图像会发生闪烁，当你观看具有大面积高度饱和颜色的画面时尤其如此。使用100Hz隔行帧率（也称为双扫描）可以使图像质量得到大幅改善。

早期的100Hz的电视只是简单地对场进行重复（$F_1F_1F_2F_2F_3F_3F_4F_4\cdots$），如图7-45a所示。然而，它们仍然会存在线闪烁，水平线总是在奇数线和偶数线之间跳跃。这种扰动每1/25秒会发生一次。

使用场序列$F_1F_2F_1F_2F_3F_4F_3F_4\cdots$可以解决线闪烁的问题。但这会引起运动画面的抖动。对静态图像，可以使用图像序列$F_1F_2F_1F_2F_3F_4F_3F_4\cdots$对其进行补偿；而对于运动图像，可以使用图像序列$F_1F_1F_2F_2F_3F_3F_4F_4\cdots$对其进行补偿。

如果观看的节目是为电影而制作的，那么仍然不能获得理想的画面。它们容易受到抖动的影

响，因为每个电影帧需要传送两次。这种情况下不采用场序列F_1F_1 F_2F_2 F_3 F_3 F_4 $F_4\cdots$，而是要采用场序列F_1 F_2F_1 F_2 F_3 F_4F_3 $F_4\cdots$（如图7-45b所示），其中F_n是F_n和F_n+1之间的一幅对运动进行补偿的生成图像。



（a）垂直低通滤波

（b）按照中间顺序格式进行重采样以及时域低通滤波

（c）按照最终的标准进行重采样

图7-42 其后紧接时域插值的垂直插值的谱表示

236

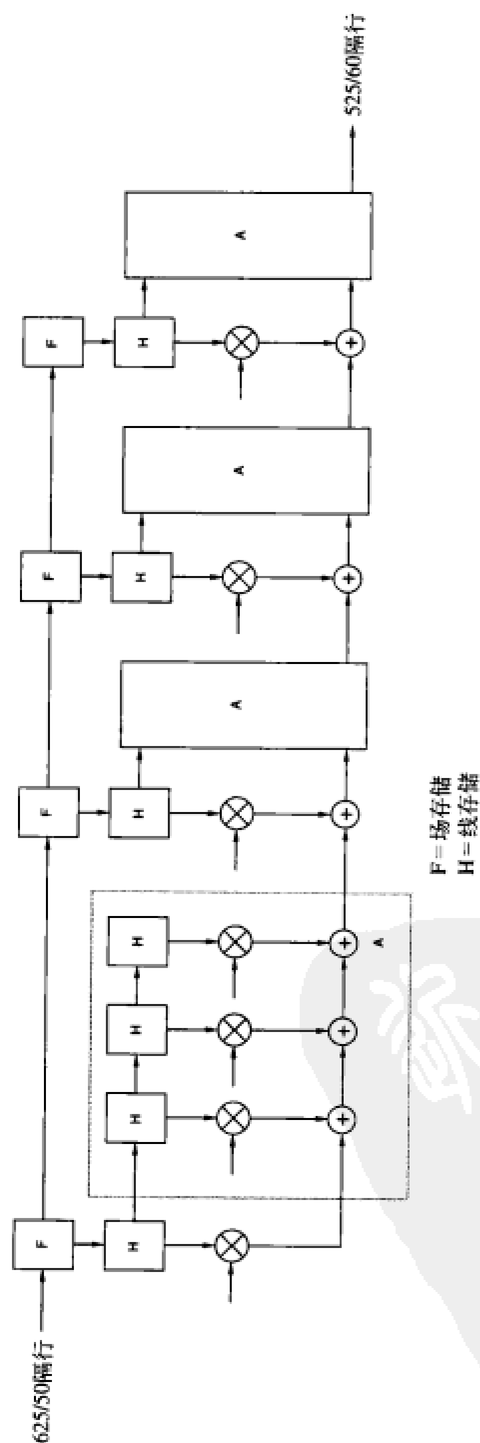


图7-43　典型的50Hz到60Hz帧频的变换，使用的是组合的垂直和时域插值

237

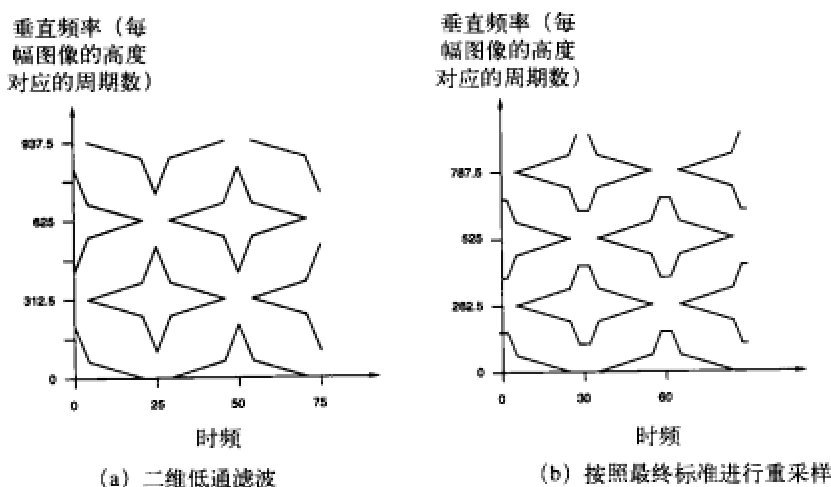


图7-44 组合的垂直和时域插值的谱表示

238

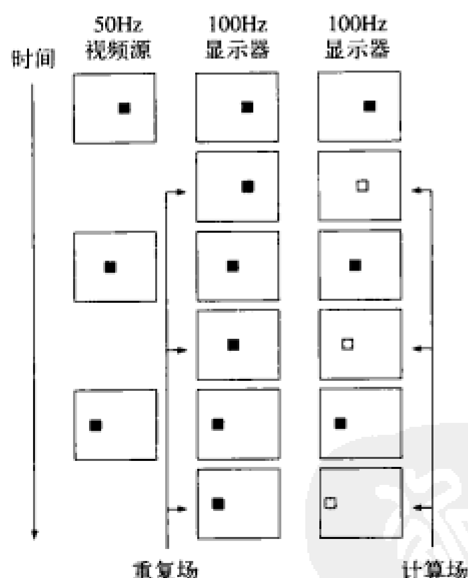


图7-45 50Hz到60Hz帧率（隔行双扫描）的转换技术

239

7.8.3 2：2下拉模式

这种技术用在50Hz帧率的某些基于电影的压缩视频内容。通常，电影是每秒钟记录24帧。

在压缩期间，电视电影机从每秒24帧加速到了每秒25帧，制作的内容为每秒25帧的逐行视频。在解压期间，每个电影帧简单地映射为两个视频场（得到的是576i25或1080i25的视频）或两个视频帧（得到的是576p50、720p50或1080p50的视频）。

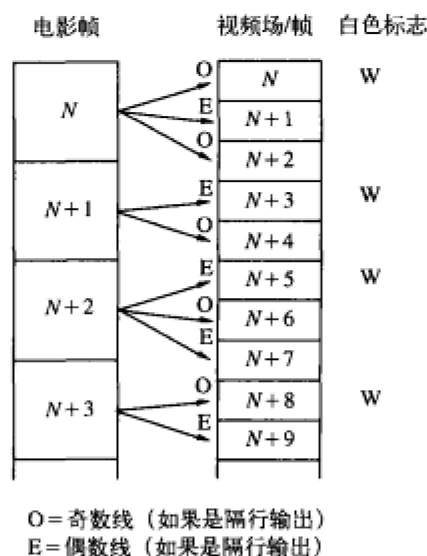
这种技术提供了更高的视频质量，并避免了运动抖动伪影。然而，这种技术将视频节目的播放时间缩短了4%，将2小时的电影截去了约5分钟。音频解码器不能够处理由S/PDIF（IEC 60958）传输的快4%的音频数据。

为了对电视电影机加速引起的音频变化量（pitch）进行补偿，在解码期间，可以对其重采样

以恢复初始的量（在低端消费类产品上是很耗时的），或在节目制作过程中可以进行重采样。

7.8.4　3：2下拉模式

当将每秒24帧的视频内容变换到60Hz帧率时，通常使用3：2下拉模式（pulldown），如图7-46所示。在压缩过程中，电影播放速度减慢了0.1%，变为每秒23.976（24/1.001）帧，因为使用59.97Hz是为了与NTSC时序兼容。在解压过程中，2个电影帧会产生5个视频场（得到的是480i30或1080i30的视频），或5个视频帧（得到的是480p60、720p60或1080p60的视频）。



240

图7-46　将24Hz的电影转换为60Hz视频的3：2下拉模式

在具有高速运动对象的场景中，对于用于特定视频场或帧的特定电影帧，可以手动进行调整以最小化运动伪影。

在视频解压的过程中，也可以使用3：2下拉模式简单地将帧率从每秒23.976（24/1.001）帧增加到59.97（60/1.001）帧，避免去隔行的问题。

可以使用变速器来掩盖某些问题，如瑕疵（defect）、拼接（splicing）和审查剪切（censorship cut），以及改变节目的播放时间。变速器不会重复电影帧，也不会引起不流畅，它只是破坏了电影和视频之间3：2的比例关系，并且所破坏视频的长度要足够长，这样才能确保平滑的时间速率。

为了获得最优的静止帧性能，模拟激光影碟使用白色标志信号来指示相关场的另一个序列的开始。在静止帧模式下，白色标志信号告诉系统备份两个场（使用它们之间没有运动的两个场）来重新显示当前帧。

7.8.5　3：3下拉模式

这种技术用在某些显示器中支持72Hz的帧率。通过简单地复制每个电影帧三次，就可以将每秒24帧基于电影的视频内容转换为72Hz的隔行视频。

7.8.6　24：1下拉模式

这种技术也称为12：1下拉模式，可以用来将24帧/秒的帧率转换为50场/秒。

除了每12个电影帧的最后一帧生成3帧视频场以外，其余每个电影帧生成两个视频场。当画面中的运动比较平滑时，尽管声调变化量（audio pitch）是正确的，但是每半秒钟还是会出现运动抖动。

7.9 逐行到隔行的转换

在有些应用中，在隔行显示器上显示逐行视频信号是必要的。这样，就需要某种形式的变换实现逐行到隔行视频的转换。

逐行到隔行变换必须在分量视频信号上进行（如R'G'B'或YCbCr），复合颜色视频信号（如NTSC或PAL）不能直接处理，这是由于复合颜色视频信号上存在彩色副载波相位信息，如果直接对其进行处理，在处理以后这些信息就变得没有意义了。在变换之前，必须将这些信号解码为分量颜色信号，如R'G'B'或YCbCr。

有2种基本的逐行到隔行转换技术：扫描线抽取（scan line decimation）和垂直滤波（vertical filtering）。

7.9.1 扫描线抽取

最简单的方法就是在每个逐行帧中，每隔一行有效扫描线就丢弃一行，如图7-47所示。尽管这种方法的代价最小，但是它也存在问题，尤其是对于位于画面顶部和底部的对象，问题更加明显。

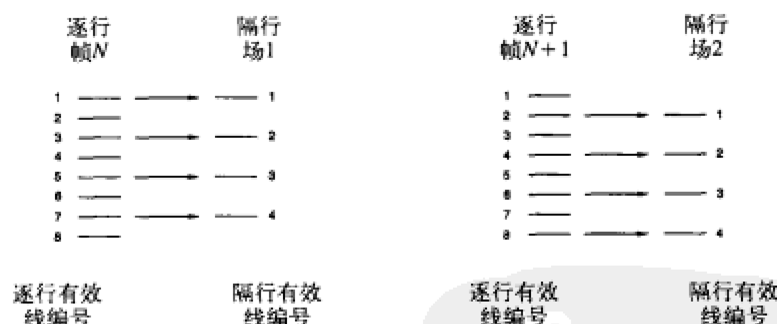


图7-47 使用扫描线抽取实现逐行到隔行的变换

如果在垂直方向上存在明显的颜色或亮度过渡，那么，仍然会以一半帧率的速度发生闪烁。其原因是由于抽取导致每隔一场才显示一场。例如，一个逐行扫描线宽度的水平线将会时隐时现，从而导致闪烁。具有两倍逐行扫描线宽度的水平线将会上下振荡。

简单的抽取操作也可能增加混叠伪影。尽管这种混叠伪影不一定可见，但它会影响到未来对图像的处理。

241
242

7.9.2 垂直滤波

一个更好的解决方案是使用两个或更多行的逐行扫描线数据来生成一行隔行扫描线数据。快速的垂直过渡可以在多个隔行线上进行平滑。

对于3线滤波器，如图7-48所示，典型的滤波系数为[0.25, 0.5, 0.25]。使用多于3线的滤波通常会导致过度的模糊，从而难以看清小的文字。

另外一种可能的实现是使用IIR滤波，而不是FIR滤波。除了平均以外，这种技术还降低了对象周围的亮度，从而更进一步减少了闪烁。

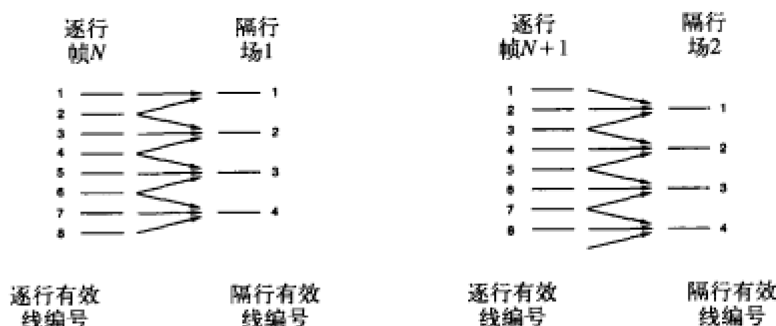


图7-48　使用3线垂直滤波实现视频从逐行到隔行的变换

注意，每帧的开始与结束处，我们必须慎重，因为用于滤波的扫描线较少。

7.10　隔行到逐行变换

在有些应用中，必须在逐行显示器上显示隔行视频信号。这样，就可能需要某种形式的*去隔行*（deinterlaced）或*逐行扫描*（progressive scan）变换。

注意，去隔行必须在分量视频信号上进行（如R'G'B'或YcbCr）。复合颜色视频信号（如NTSC或PAL）不能直接进行去隔行处理，因为复合视频信号上存在颜色副载波相位信息，如果直接进行处理，那么处理之后它就变得没有意义了。在去隔行处理之前，必须将这些信号解码为分量颜色信号，如R'G'B'或YCbCr。

有两种基本的去隔行算法：视频模式和电影模式。视频模式去隔行可以进一步分解为场内处理和场间处理。一个良好的去隔行器的目标是在特定的时刻正确地选择所需的最好算法。

在视频源和显示器的垂直分辨率不匹配的系统中（例如，由于在HDTV上显示SDTV视频内容而导致的不匹配），去隔行处理和垂直缩放处理可以合并为一个单一的处理过程。

7.10.1　视频模式：场内处理

这是仅使用初始场中的信息生成额外扫描线最简单的方法。计算机行业已经为这种技术创造了一个新的词*单场插值*（bob）。

尽管实现场内处理有2种常用技术——扫描线复制和扫描线插值，但是最后的垂直分辨率总是受限于初始场的内容。

1. 扫描线复制

扫描线复制（scan line duplication）（见图7-49）简单地复制以前的有效扫描线。尽管有效扫描线的数目加倍了，但垂直分辨率并没有增加。

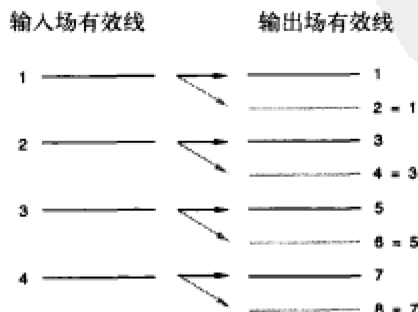


图7-49　使用扫描线复制实现去隔行。通过复制新扫描线以上的有效扫描线来生成新扫描线

2. 扫描线插值

扫描线插值在初始的有效扫描线之间生成插值的扫描线。尽管有效扫描线的数目加倍了，但垂直分辨率并没有加倍。

如图7-50所示，最简单的实现是使用线性插值在两个输入扫描线之间生成一行新的扫描线：

$$out_n = (in_{n-1} + in_{n+1})/2$$

通过使用FIR滤波器，可以获得更好的结果，但是需要付出额外的代价：

243

$$\begin{aligned} out_n = {} & 160(in_{n-1} + in_{n+1}) \\ & - 48(in_{n-3} + in_{n+3}) \\ & + 24(in_{n-5} + in_{n+5}) \\ & - 12(in_{n-7} + in_{n+7}) \\ & + 6(in_{n-9} + in_{n+9}) \\ & - 2(in_{n-11} + in_{n+11}) \end{aligned}$$

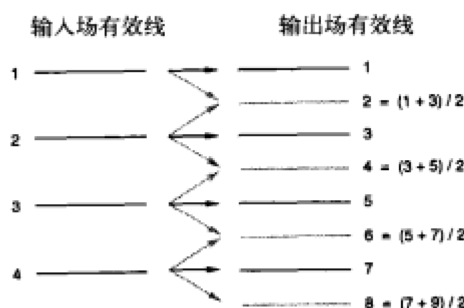


图7-50 使用扫描线插值实现去隔行。通过对前一和下一有效扫描线取平均来生成新扫描线

3. 分数比率插值

在很多情况下，在输入扫描线与输出扫描线的数目之间存在周期性的关系，但这种周期不是整数。在这种情况下，就可能需要进行分数比率插值，这种插值类似于仅在垂直方向上进行缩放的多相位滤波（polyphrase filtering）。这种技术将去隔行和垂直缩放处理结合起来形成一个处理过程。

4. 可变插值

在少数情况下，输入与输出扫描线编号之间并不存在周期性关系。所以，在理论上，需要无限多的滤波器相位和系数。因为这是不可行的，所以解决方案是采用大量的但有限的滤波器相位。滤波器相位的数量决定了插值准确度。这种技术也将去隔行和垂直缩放处理结合起来形成一个处理过程。

7.10.2 视频模式：场间处理

在这种方法中，需要使用来自多个（一个以上）场的视频信息生成一个单一的逐行帧。因为它使用了来源于多个场的视频信息，所以这种方法可以得到更高的垂直分辨率。

1. 场合并

这种技术将两个连续场合并，生成一个视频帧（如图7-51所示）。在每个场时间内，这个场的有效扫描线与前一场的有效扫描线合并。其结果是，对于每个输入场时间，结合一对场生成一帧（见图7-52）。尽管这种方法实现简单，但是只有在没有运动的区域中垂直分辨率才加倍。

244 ~ 245

运动对象将出现伪影，也称为梳形，这是由于两个场之间存在时差造成的：从一个场到另一个场，运动对象定位在不同的位置。当将两个场合并以后，运动对象将出现一幅画面重叠的图像（如图7-53所示）。

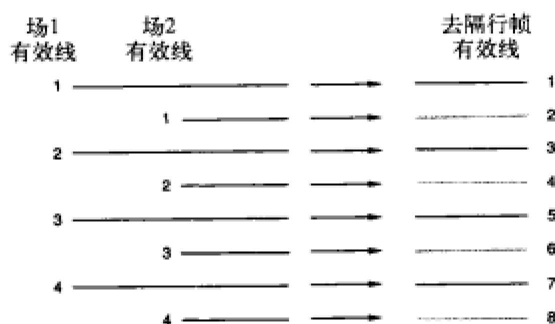


图7-51　使用场合并方法实现去隔行。通过使用下一场或前一场的扫描线作为输入生成阴影扫描线

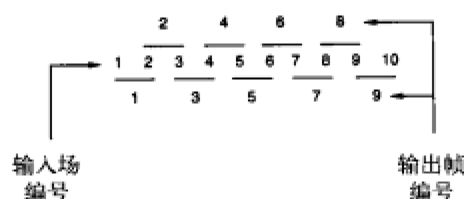


图7-52　以场率生成去隔行帧

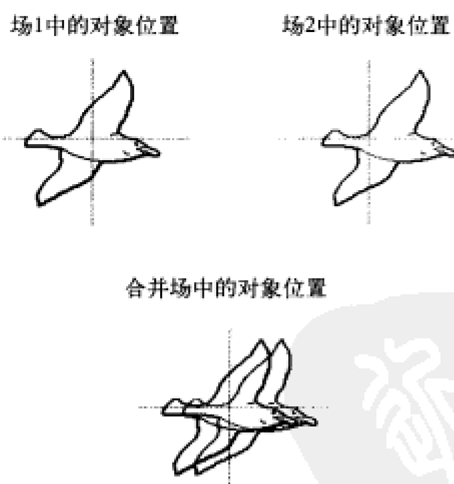


图7-53　当使用场合并时产生的运动伪影

通常在垂直方向上对图像稍微进行软化，试图降低伪影的可见度。当实现垂直方向上的软化时，会引起垂直分辨率的损失，也会导致运动和平移抖动。

计算机行业将这种技术称为场合并（weave），但是编织也包含逆向电视电影过程以消除视频源中出现的任何3：2下拉。从理论上讲，这消除了双重图像伪影，因为现在合并的是两个相同的场。

2. 运动自适应去隔行

一个优秀的去隔行解决方案是，在图像的静止画面区域使用场合并，而在动态画面区域使用扫描线插值。为了以逐样本点的方式实现这种方案，必须在整个画面中实时地检测运动区域和静止区域，这要求对视频的多个场进行处理。

当两个场合并时，在图像的静止区域中仍然完全维持了原有的垂直分辨率，在静止区域中，眼睛对细节是最敏感的。样本点的差异值可以达到任何值，从0（没有运动和噪声）到最大值（例如，从最亮的颜色变到黑色）。我们必须选择什么时候使用来自前一场的样本点（由于运动使得其处在错误的位置上），或什么时候由当前场的相邻扫描线插值生成新样本点。两种方法之间的突然切换在视觉上是可以觉察到的，因此，需要使用交叉过渡（也称为软切换）。当样本差异处于某些幅度时，双重图像导致的分辨率损失与插值导致的分辨率损失是相当的。运动量应该使得交叉过渡器处于50%的位置。过少的运动将导致过渡倾向场合并，而过多的运动将导致过渡倾向插值结果。

不采用"每像素"的运动自适应去隔行，因为这需要对每个像素进行决策。某些低代价的解决方案使用"每场"的运动自适应去隔行。在这种情况下，基于场之间的运动量针对每个场选择算法。尽管"每像素"的运动自适应去隔行实现困难，但是如果处理得当，画面看起来非常好。"每场"的运动自适应去隔行看起来不比垂直插值好多少。

3. 运动补偿去隔行

运动补偿（motion compensated）（也称为运动向量操作）去隔行比运动自适应去隔行的复杂程度要高几个数量级，通常见于视频格式转换器中。

运动补偿处理要求在场之间为每个样本点计算运动向量，沿着每个样本点的运动轨迹进行插值。也必须找到通过每个丢失样本点的运动向量。当你在帧之间移动时，图像的某些区域可能被遮挡，或露出来。运动向量也必须到达亚像素的准确度，并且在帧之间的两个时间方向上确定运动向量。

MPEG使用的运动向量误差是自校正的，因为对预测宏块之间的残留差值也进行了编码。当运动补偿去隔行是单个终端系统时，运动向量误差将产生伪影，因此必须使用不同的搜索和确认算法。 246

7.10.3 电影模式（使用逆向电视电影）

对于具有3：2下拉模式的视频源（也就是说，从24帧/秒的电影转换为60场/秒的视频），在处理之前消除复制场可以获得更高的去隔行性能。

逆向电视化过程对3：2的场序列进行检测，将冗余的第3场去掉。将剩下的场对合并（因为它们之间不存在运动）以形成24帧/秒的逐行视频帧。然后在3：2序列中重复这些帧以得到60帧/秒的帧率。

尽管看起来这好像是理想的解决方案，但是在同一个节目中，某些视频内容同时使用了60场/秒的视频和24帧/秒的视频（基于电影）。另外，某些视频内容可能偶尔同时出现了两种类型的视频。在另外一种情况中，3：2下拉时序不能保持规律性，或者，视频源开始并不是来自于电影。这样，去隔行器必须对每种视频类型进行检测，并采取不同的处理模式（视频模式与电影模式）。通常会出现显示器伪影，这是由两种视频类型变换之间的延迟和去隔行器对这种变化的检测引起的。

7.10.4 关于频率响应的注意事项

用于去隔行的各种两倍垂直向上采样（upsampling）技术的实现：在两行有效线之间填充0值，然后进行滤波，如图7-54所示。

线A给出了线复制的频率响应，在线A中，对于所给滤波器的低通滤波器系数为1、1和0。

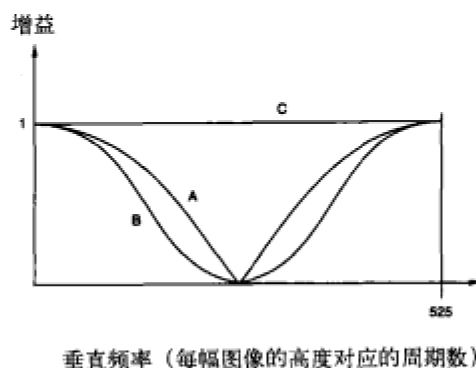


图7-54　各种去隔行滤波器的频率响应。A为线复制，B为线插值，C为场合并

使用低通滤波器系数0.5、1.0和0.5的线插值得到的频率响应曲线为线B。注意，线复制将得到更好的高频响应。具有比线复制更好频率响应的垂直滤波器是可能实现的，只是以更多的线存储和更多的处理为代价。

247

在实现了场合并后，可以获得最好的垂直频率响应。线的空间位置已经是正确的，并且不需要垂直处理，这样可以得到更平坦的曲线（曲线C）。同样，这只是用于图像的静态区域。

7.11　基于DCT的压缩

很多视频压缩的变换过程是建立在离散余弦变换（Discrete Cosine Transform，DCT）的基础上的。要想象DCT是怎么回事，最容易的方式是将其当作一个滤波器组（filter bank），这些滤波器都是并行计算的。

在编码期间，通常在DCT之后还要进行几个其他的操作，如量化、Z形扫描（zig-zag scanning）、行程长度编码和可变长编码。在解码期间，这个处理过程刚好相反，并且其中的每个处理都是其对应于编码的逆过程。

当讨论视频压缩的时候，我们多次使用了术语宏块（macroblock）和块（block）。图7-55给出了这两个术语之间的关系，并指明了为什么变换处理通常是以8×8的样本块为单位。MPEG-4.10（H.264）也支持8×4、4×8和4×4的块。

7.11.1　DCT

8×8 DCT对一个8×8的样本块进行处理，生成一个8×8的DCT系数块（如图7-56所示）。DCT变换的输入可以来自实践的视频帧或运动补偿微分（误差）值，这取决于编码器的操作模式。每个DCT系数表示的是块内特定的水平或垂直频率的量。

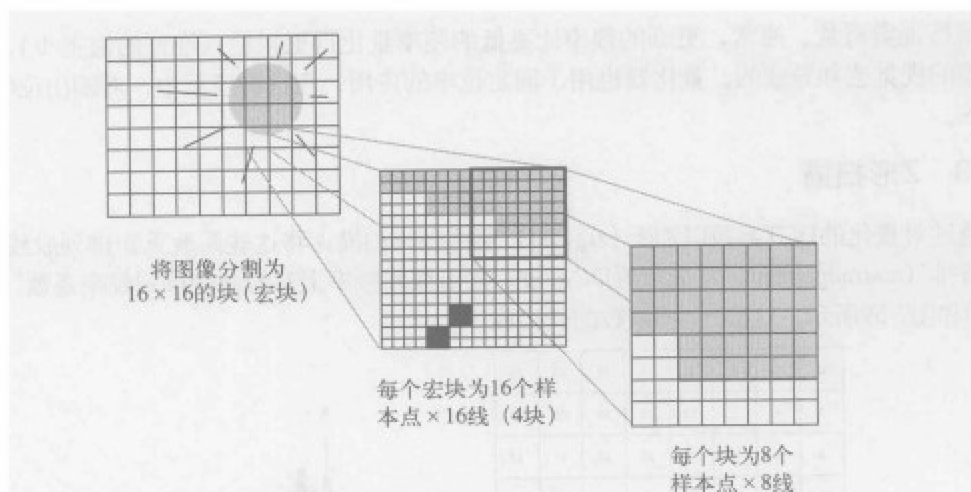


图7-55 宏块与块之间的关系

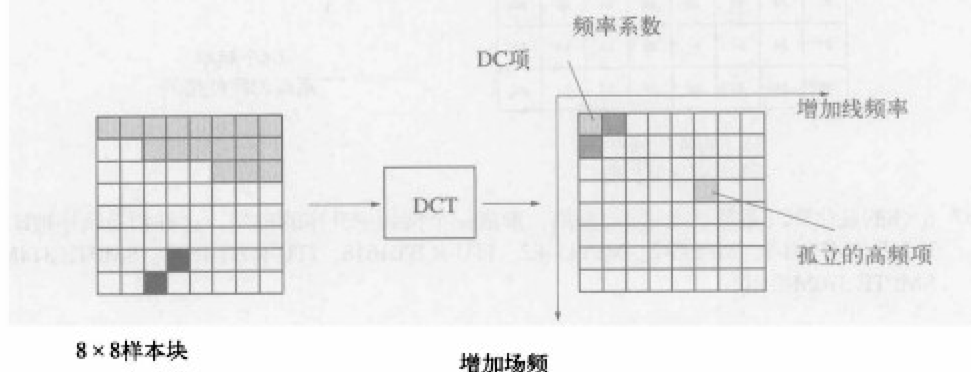


图7-56 使用DCT处理8×8样本块或残差项以生成8×8 DCT系数块

DCT系数（0,0）是DC系数或平均样本值。由于自然图像中样本与样本之间倾向于只会发生细微的变化，因此典型情况下，低频系数的值较大，而高频系数的值较小。

重建的8×8样本块是使用8×8逆DCT（IDCT）来生成的。尽管在理论上可以实现准确重建，但是由于有限精度的数学运算、量化和不同的IDCT实现的缘故，准确重建是不现实的。其结果是，不同的IDCT实现之间存在失配现象。

通过消除对失配具有最大贡献的位模板，失配控制试图降低编码器与解码器IDCT结果之间的漂移（drift）。

MPEG-1失配控制就是我们所知的“奇异化（oddification）”，这是由于它强制所有的量化DCT系数为负值。MPEG-2和MPEG-4.2使用了一个被称为“LSB捆绑”的改进方法，在逆向量化后，它只影响第63个DCT系数的LSB。

H.264（也就是我们所知的MPEG-4.10）通过使用“准确匹配逆变换（exact-match inverse transform）”完美地回避了这个问题。每个解码器都将准确地生成同样的图像，其他所有的方面也是相同的。

248
~
249

7.11.2 量化

对8×8的DCT系数块进行了量化，这降低了整数系数的整体精度，并有可能消除高频系数，

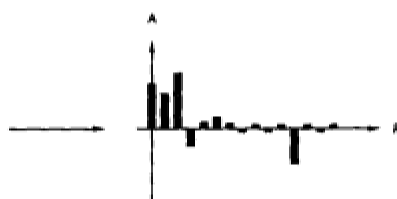
同时保持视觉质量。通常，更高的频率比更低的频率量化得更粗糙（准许的值更少），这是由量化误差的视觉感知导致的。量化器也用于固定位率的应用，在这种应用中，对输出位率的控制是变化的。

7.11.3 Z形扫描

通过对量化的DCT系数以Z形（zig-zag）顺序进行扫描，将这些系数重新排列成线性数据流。这种重排（rearrangement）首先放置DC系数，接着按照频率上升的顺序排列频率系数，如图7-57、图7-58和图7-59所示。这会产生0系数的长行程。

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

8×8的量化频率
系数块的Z形扫描

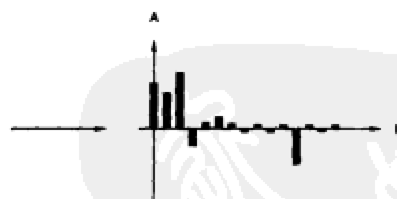


64个频率
系数的线性排列

图7-57 8×8的量化DCT系数块通过Z形扫描，形成一个按频率升序的排列。这种扫描顺序被H.261、H.263、MPEG-1、MPEG-2、MPEG-4.2、ITU-R BT.1618、ITU-R BT.1620、SMPTE 314M和SMPTE 370M采用

0	4	6	20	22	36	38	52
1	5	7	21	23	37	39	53
2	8	19	24	34	40	50	54
3	9	18	25	35	41	51	55
10	17	26	30	42	46	56	60
11	16	27	31	43	47	57	61
12	15	28	32	44	48	58	62
13	14	29	33	45	49	59	63

8×8的量化频率系
数的Z形扫描



64个频率系数
的线性排列

图7-58 H.263、MPEG-2和MPEG-4.2 “交替-垂直（alternate-vertical）” 扫描排序

7.11.4 行程长度编码

将量化的频率系数线性流变换为一系列的[run, amplitude]对。[run]表示系数0的个数，而[amplitude]表示紧接在run后的非0系数。

7.11.5 可变长编码

使用可变长码对[run, amplitude]对进行编码，得到的是另一个无损压缩。对于公共数据对而

言，这种压缩方法可以得到较短的编码，而公共数据对较少时，则可能得到较长的编码。

这种编码得到的是DCT系数的一个更加紧凑的形式，因为大量的DCT系数通过量化以后变为了0，重排序以后，可以得到一组长游程的连续0值。 250↓251

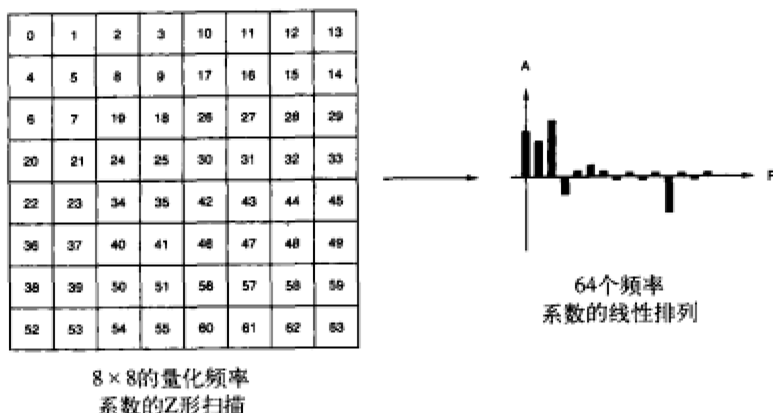


图7-59 H.263和MPEG-4.2“交替－水平（alternate-horizental）”扫描排序

7.12 关于固定像素显示器的注意事项

固定像素显示器独特的设计和颜色再生域促成了多种新视频处理技术的发展。这使得不管视频源是什么样的，都可以得到亮度更高、对象边缘更明显和颜色更丰富的图像。

7.12.1 扩展颜色再生

通常需要对广播站进行调整以满足基于CRT的电视机有限的颜色再生特性。为了适应PDP和LCD的颜色再生能力，制造商已经引入了各种颜色扩展（color expansion）技术。这些颜色扩展包括针对每个基本色或辅助色以及肤色（flesh color）使用独立的色度和饱和度控制。

7.12.2 细节校正

在CRT电视机中，增强图像通常是通过改变电子束的直径来完成的。对于固定像素的显示器，在视频信号上增加超调（overshoot）和负调（undershoot）将导致失真。一个可以接受的实现是，在需要轮廓增强的区域之间和之后，逐渐改变图像的亮度。 252

7.12.3 非均匀量化

不是简单地增加量化等级，而是让量化步长可以实现与图像的亮度一致的变化。这是可能的，因为，相对于亮图像中的亮度变化而言，人更容易检测到暗图像中亮度的小变化。另外，图像越亮，人对亮度的变化越不敏感。这意味着，对于暗图像可以使用比亮图像更多的量化等级。这种技术也可以用于增加图像中频繁出现的阴影的量化等级。

7.12.4 缩放与去隔行

在去隔行期间（是为了让折叠失真引起的波纹噪声最小化），固定像素显示器，如LCD和等离子显示器，通常先进行放大然后再缩小。例如，将一个1080i视频源去隔行得到2160p的视频，然后再缩放到1536p，并最终缩放到768p（以驱动1024×768的显示器）。作为选择，某些解决方

案将视频去隔行并缩放到1500p，然后再缩放到显示器的本身分辨率。

7.13　应用实例

图7-60和图7-61给出了在视频解压和去隔行后的典型视频处理。

除了基本视频源外，其他的典型视频源包含屏幕控制系统（On-Screen-Display，OSD）、内容导航图形、隐藏字幕（closed captioning）或字幕（subtitle）和画中画（picture-in-picture，PIP）的第二视频。

OSD平面（plane）显示配置菜单盒，如视频输出格式和分辨率、音频输出格式等。对于每种产品，OSD设计都是不同的，因此OSD平面通常支持范围广泛的RGB/YCbCr格式和分辨率。查询表gamma校正线性RGB数据，将2、4或8位的索引颜色转换为32位的YcbCrA数据，或将0~255的图形颜色等级转换到16~235的视频颜色等级。

内容导航平面（content navigation plane）显示由蓝光BD-J、HD DVD HDi、电子节目向导等生成的图形。它应该支持与OSD平面相同的格式和能力。

字幕平面（subtitle plane）是一个有用的区域，用于渲染字幕信息、DVB字幕、DVD子画面（subpicture）等。查询表将2、4或8位索引颜色转换为32位的YcbCrA数据。

辅助视频平面（secondary video plane）通常用于支持第二个视频源以实现画中画（PIP）或图形显示（如JPEG图片）。对于图形数据，查询表可以对线性RGB数据进行gamma校正，将2、4和8位的索引颜色转换为32位的YCbCrA数据，或将0~255的图形颜色等级转换为16~235的视频颜色等级。

能够独立地对每个视频源进行缩放，就可以提供最大的灵活性。除了不论视频源的分辨率是多少都能够输出任意的分辨率外，这样也可以包容一些特效。

色度校正确保了颜色是准确的，而且与视频源和显示器没有关系（SDTV与HDTV）。

对每个视频源和视频输出接口提供独立的亮度、对比度、饱和度、色度和锐度控制，这样可以具有最大的灵活性。例如，可以对PIP进行调整而不影响主要画面，可以对视频进行调整而不会影响静态画面视频质量等。

在图7-61中顶部的NTSC/PAL编码器中，由于使用了可选的缩小变换和隔行－逐行变换模块，因此可以同时输出HD和SD，或同时输出逐行和隔行视频，而不会影响HD或逐行视频质量。

当摄录过程不会出现偶尔记录OSD或字幕信息时，图7-61底部显示的第二个NTSC/PAL编码器是有用的。

253

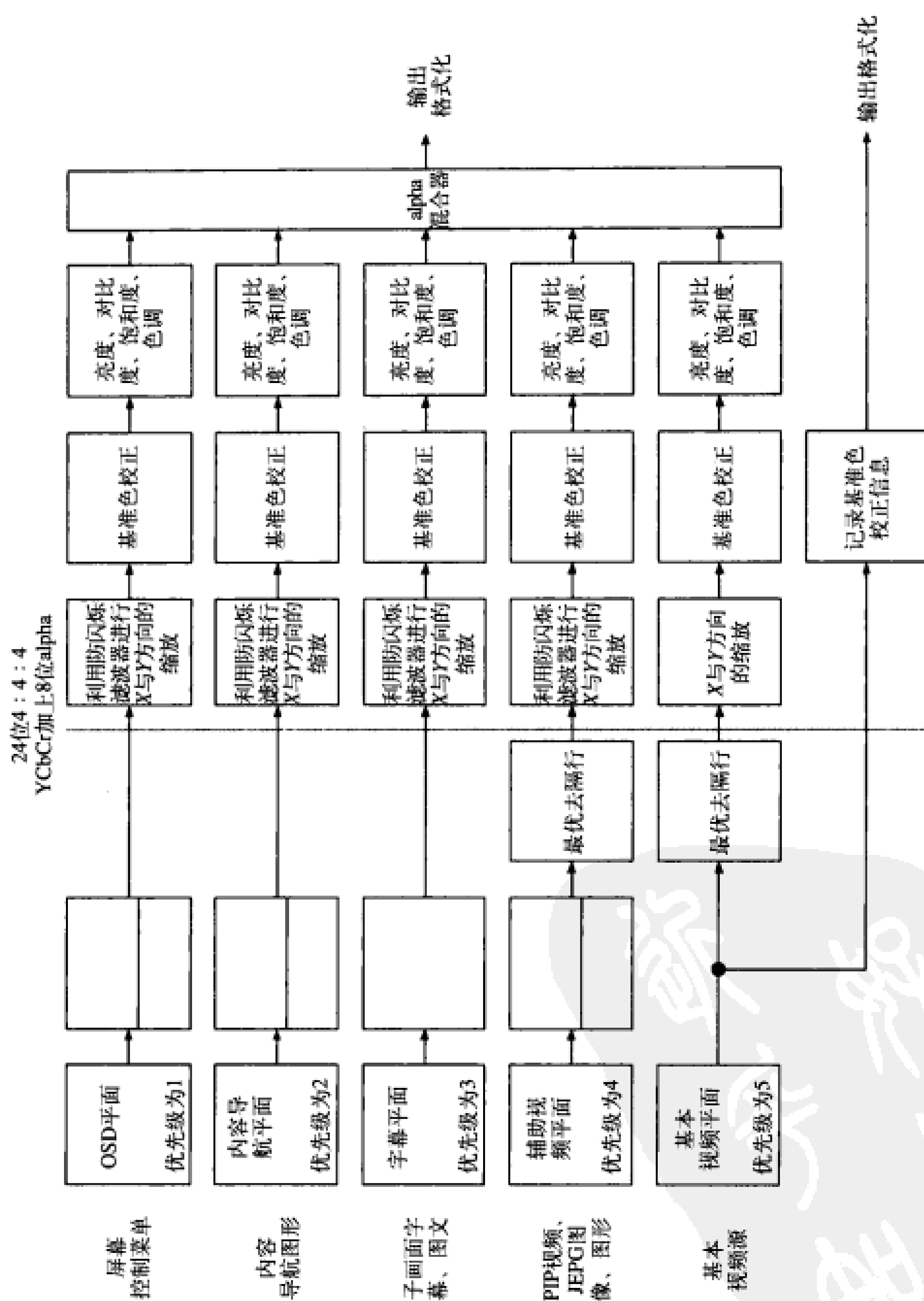
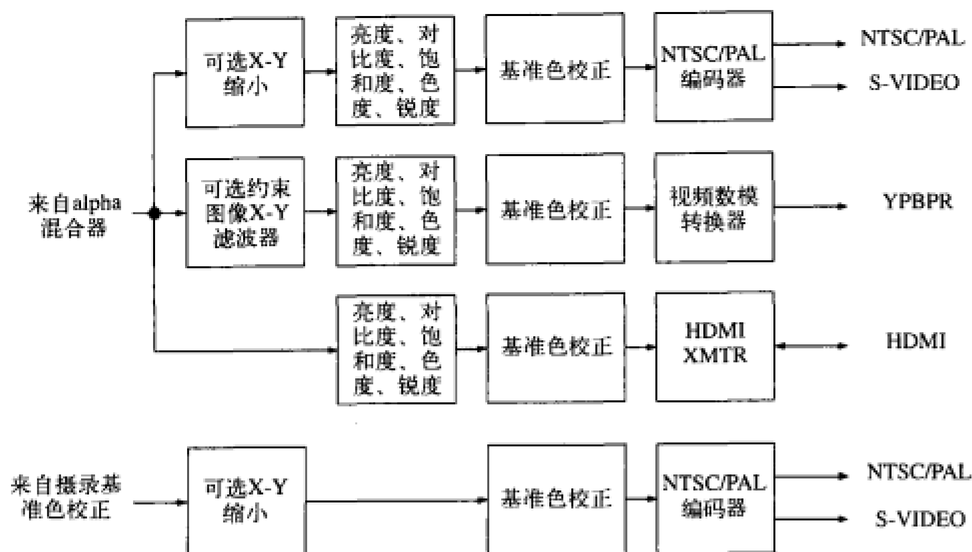


图7-60 视频复合的简化框图



255

图7-61 视频输出端口处理

参考文献

1. Clarke, C.K.P., 1989, *Digital Video: Studio Signal Processing*, BBC Research Department Report BBCRD 1989/14.
2. Devereux, V. G., 1984, *Filtering of the Colour-Difference Signals in 4 : 2 : 2 YUV Digital Video Coding Systems*, BBC Research Department Report BBC RD1984/4.
3. ITU-R BT.601-5, 1995, *Studio Encoding Parameters of Digital Television for Stantard 4 : 3 and Widescreen 16 : 9 Aspect Ratios.*
4. ITU-R BT.709-5, 2002, *Parameter Values for the HDTV Standards for Production and International Programme Exchange.*
5. ITU-R BT.1358, 1998, *Studio Parameters of 625 and 525 Line Progressive Scan Television Systems.*
6. Johan G.W.M. Janssen, Jeroen H. Stessen, and Peter H.N. de With, *An Advanced Sampling Rate Conversion Technique for Video and Graphics Signals, Philips Research Labs.*
7. Sandbank, C. P., *Digital Television*, John Wiley&Sons,Ltd., New York, 1990.
8. SMPTE 274M-2005, *Television——19201 × 080 Image Sample Structure, Digital Timing Reference Sequences for Multiple Picture Rates.*
9. SMPTE 293M-2003, *Television——720 × 483 Active Line at 59.94Hz Progressive Scan Production——Digital Representation.*
10. SMPTE 296M-2001, *Television——1280 × 720 Progressive Image Sample Structure, Analog and Digital Representation and Analog Interface.*
11. SMPTE EG36-1999, *Transformations Between Television Component Color Signals.*
12. Thomas, G. A., 1996, *A Comparison of Motion-Compensated Interlace-to-Progressive Conversion Methods*, BBC Research Department Report BBC RD1996/9.
13. Ultimatte®, Technical Bulletin NO.5, Ultimatte Corporation.
14. Watkinson, John, *The Engineer's Guide to Standards Conversion*, Snell and Wilcox Handbook Series.
15. Watkinson, John, *The Engineer's Guide to Motion Compensation*, Snell and Wilcox Handbook Series.

256

第8章　NTSC、PAL和SECAM概述

为了全面理解NTSC、PAL和SECAM的编码和解码过程，有必要先回顾一下这些标准的背景以及它们的产生过程。

8.1　NTSC概述

世界上第一种彩色电视制式是在美国开发的。美国联邦通信委员会（Federal Communications Commission，FCC）于1953年12月17日批准了这一传输标准，并批准从1954年1月23日开始广播。在开发彩色电视传输标准的过程中，美国国家电视制式委员会（National Television System Committee，NTSC）所做的大部分工作就是使之与（当时的）525行、60场/秒、2：1隔行黑白电视标准兼容。

8.1.1　亮度信息

黑白亮度（Y）信号由gamma校正后的红绿蓝（R′G′B′）信号表示：

$$Y=0.299R'+0.587G'+0.114B'$$

由于伴音副载波位于4.5MHz，所以要求彩色信号位于和黑白视频信号相同的带宽（0~4.2MHz）之内。

基于经济上的考虑，NTSC还有另外一个要求，即黑白电视接收机必须能显示彩色电视广播的黑白部分，而彩色电视接收机也必须能显示黑白电视广播。

8.1.2　彩色信息

眼睛对亮度的空间和时间变化最为敏感，所以仍然允许亮度信息覆盖全部的可用带宽（0~4.2MHz）。彩色信息由色调和饱和度信息表示，眼睛对彩色信息没有亮度信息那么敏感，所以彩色信息需要较少的带宽。

色调和饱和度信息使用3.58MHz副载波传送、编码，以便接收机能够分离出色调、饱和度和亮度信息，并将它们转换回RGB信号来显示。尽管这样允许在和黑白电视信号相同的带宽内传送彩色信号，但是由于它们占用相同的频谱空间，所以如何经济地分离彩色信息和亮度信息这一难题仍然存在。

为了传送彩色信息，我们使用了“色差”信号U和V或者I和Q：

$$R'-Y=0.701R'-0.587G'-0.114B'$$
$$B'-Y=-0.299R'-0.587G'+0.886B'$$
$$U=0.492(B'-Y)$$
$$V=0.877(R'-Y)$$
$$\begin{aligned}I&=0.596R'-0.275G'-0.321B'\\&=V\cos33^\circ-U\sin33^\circ\\&=0.736(R'-Y)-0.268(B'-Y)\end{aligned}$$
$$\begin{aligned}Q&=0.212R'-0.523G'+0.311B'\\&=V\sin33^\circ+U\cos33^\circ\\&=0.478(R'-Y)+0.413(B'-Y)\end{aligned}$$

为了避免传输过程中的过调制，在由（B′−Y）和（R′−Y）导出U和V的过程中使用了压缩系数。如果使用满量程的（B′−Y）和（R′−Y），那么调制后的色度信号电平将超出黑白电视发射机所支持的电平。实验证明，调制后的副载波幅度不超出Y信号幅度的20%（黑色电平以下，白色电平以上）是允许的。选择使用一个压缩系数，从而可以使75%彩条信号的最大电平位于白色电平位置上。

最初选择使用的是I和Q，这是因为它们相对于U和V，与色敏锐度的变化更为紧密。随着所观察物体尺寸的减小，人眼对彩色的感知也逐渐减小。小物体（覆盖的频率范围为1.3~2.0MHz），人们对它的彩色感知度很低。中等大小的物体（覆盖的频率范围为0.6~1.3MHz），如果沿着橙－青色轴再生，那么还是可以接受的。更大的物体（覆盖的频率范围为0~0.6MHz）就需要再生全部三种颜色。

据此选择使用了I和Q带宽，并且通过旋转U和V色轴33°得到首选的彩色再生色轴。表示绿－紫色轴的Q分量，带宽限制在大约0.6MHz以内。代表橙－青色轴的I分量，带宽限制在大约1.3MHz以内。

将I和Q带宽分别限制在1.3MHz和0.6MHz以内的另一个优点，就是当对复合视频信号进行4.2MHz低通滤波产生不对称边带时，有利于减小这种不对称边带所产生的串扰。Q是一个双边带信号；而I则是不对称边带信号，这将有可能在I和Q之间产生串扰。Q为对称边带信号，所以不会对I产生串扰；因为Q信号带宽限制在0.6MHz以内，所以I串扰落在Q带宽之外。

现在，通常使用带宽为1.3MHz的U和V信号，而不太使用I和Q信号。在广播过程中，0.6MHz以上的UV会产生串扰，然而，由于消费设备的NTSC解码器使用的是有限UV带宽，这种串扰通常是看不见的。

图8-1示范了UV和IQ信号的矢量图。

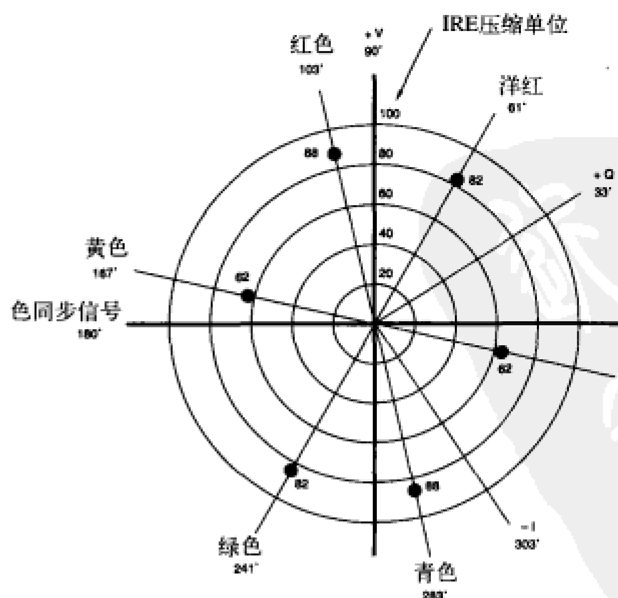


图8-1　75%彩条的UV和IQ信号矢量图

8.1.3　彩色调制

I和Q（或者U和V）使用两个相位正交的平衡调制器来调制3.58MHz的彩色副载波：其中一

个调制器由正弦相位的副载波驱动，另一个调制器由余弦相位的副载波驱动。两个调制器的输出叠加在一起形成色度调制信号： 258

$$C = Q\sin(\omega t + 33^\circ) + I\cos(\omega t + 33^\circ)$$

$$\omega = 2\pi F_{SC}$$

$$F_{SC} = 3.579545\ \text{MHz}(\pm 10\text{Hz})$$

如果使用U和V取代I和Q，那么：

$$C = U\sin\omega t + V\cos\omega t$$

与副载波相关的色度相位携带着色调信息。色度幅度携带的是饱和度信息。另外，如果物体没有颜色（例如白色、灰色或者黑色物体），那么副载波将被抑制。

8.1.4 复合视频的生成

调制色度信息加上亮度信息以及合适的水平和场同步信号、消隐信息以及色同步信号信息就生成复合彩色视频波形，如图8-2所示。

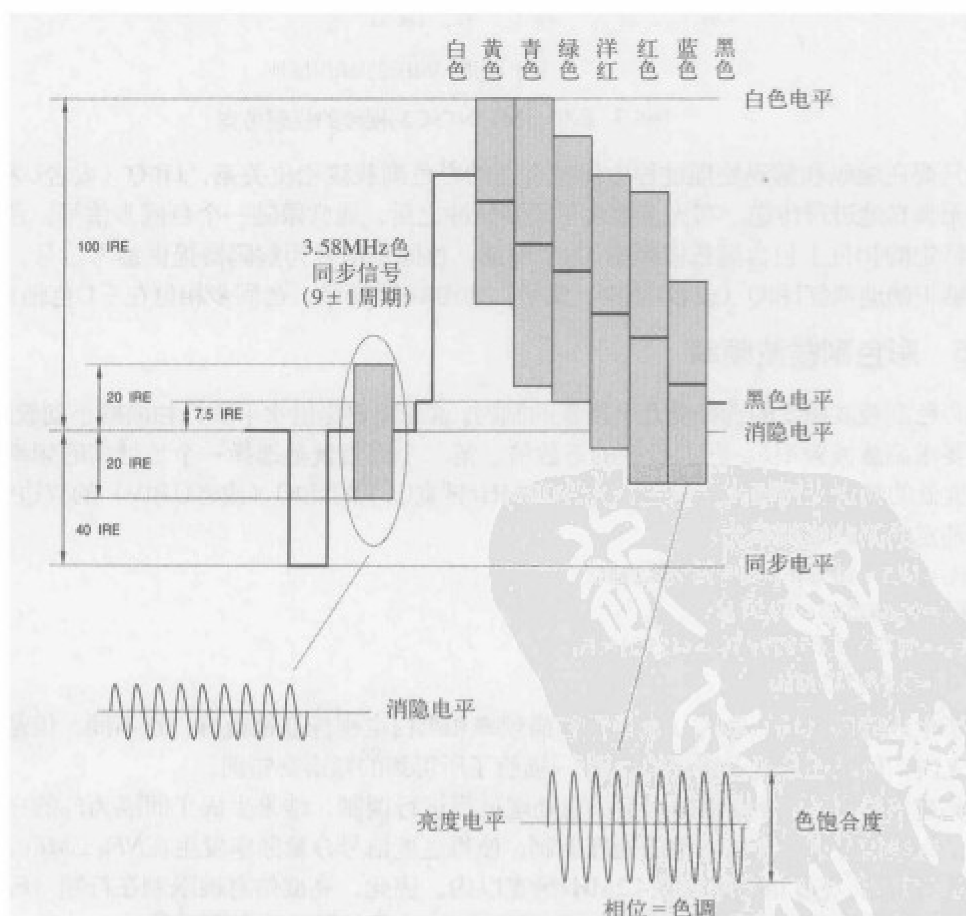


图8-2 75%彩条的（M）NTSC制复合视频信号

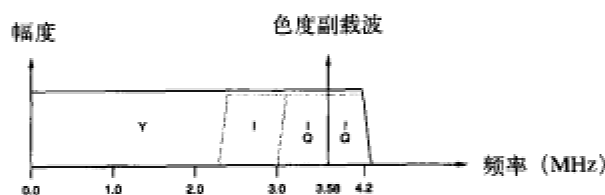
$$\text{复合NTSC} = Y + Q\sin(\omega t + 33^\circ) + I\cos(\omega t + 33^\circ) + \text{时序}$$

如果使用U和V取代I和Q，那么：

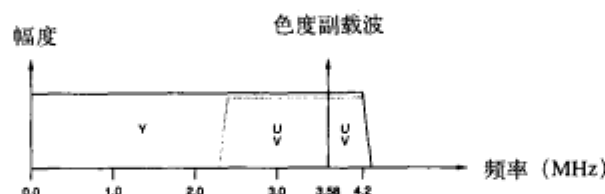
$$\text{复合NTSC} = Y + U\sin\omega t + V\cos\omega t + \text{时序}$$

259
~
260

所生成的复合视频信号的带宽如图8-3所示。



（a）使用1.3MHz的I和0.6MHz的Q信号



（b）使用1.3MHz的U和V信号

图8-3　基带（M）NTSC制视频的视频带宽

只要在编码和解码处理过程中保持合适的彩色副载波相位关系，I和Q（或者U和V）信息就可以无失真地进行传送。在大多数水平同步脉冲之后，通常跟随一个色同步信号，色同步信号在一个特定的相位上包含副载波频率的9个周期，色同步信号为解码器提供参考信号，从而使解码器能够正确地恢复I和Q（或者U和V）信号。如图8-1中所示，色同步相位在－U色轴上定义。

8.1.5　彩色副载波频率

彩色副载波频率的选择受几个因素的制约。首先需要提供水平隔行扫描减小副载波的可视度，这就要求副载波频率F_{SC}是1/2行频的奇数倍。第二个因素就是选择一个足够高的频率，生成一个可视度低的频谱交错模式。第三，允许0.6MHz带宽以下的I和Q（或者U和V）的双边带。

选定的副载波频率为：

$F_H = (4.5 \times 10^6/286)\ \text{Hz} = 15\ 734.27\ \text{Hz}$

$F_V = F_H/(525/2) = 59.94\ \text{Hz}$

$F_{SC} = ((13 \times 7 \times 5)/2) \times F_H = (455/2) \times F_H$

$\quad = 3.579545\ \text{MHz}$

所生成的F_V（场）和F_H（水平）扫描频率和黑白电视标准稍微有一点不同，但是位于容忍的范围之内，所以还是可以接受的。图8-4描绘了所得到的频谱交错图。

亮度信号（Y）分量根据水平消隐处理过程进行调制，结果生成了间隔为F_H的一束束的亮度信息。这些信号进一步被场消隐处理调制，使得亮度信号分量频率发生在$NF_H \pm MF_V$。N的最大值大约是277，将亮度信号限制在4.2MHz带宽以内。因此，亮度信息被限制在行频（F_H）整数倍谐波范围区域内，每组谐波范围内都有距NF_H29.97Hz水平帧频的附加频谱线。

两组亮度频谱之间、频率处在1/2行频奇数倍处的频谱能量最小，所以用来传送色度信息。由于彩色副载波的谐频是（1/2）F_H的奇数倍，所以它们之间相距F_H，它和亮度信号有1/2行频的偏置，结果形成了一种上移的频谱交错模式。重复一个特定的采样点位置需要4个完整的场，如图8-5所示。

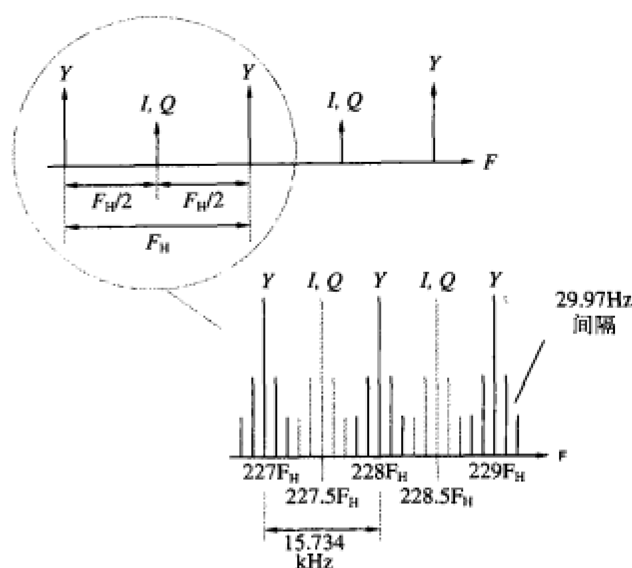


图8-4 亮度和色度频谱交错原理。注意：$227.5F_H = F_{SC}$

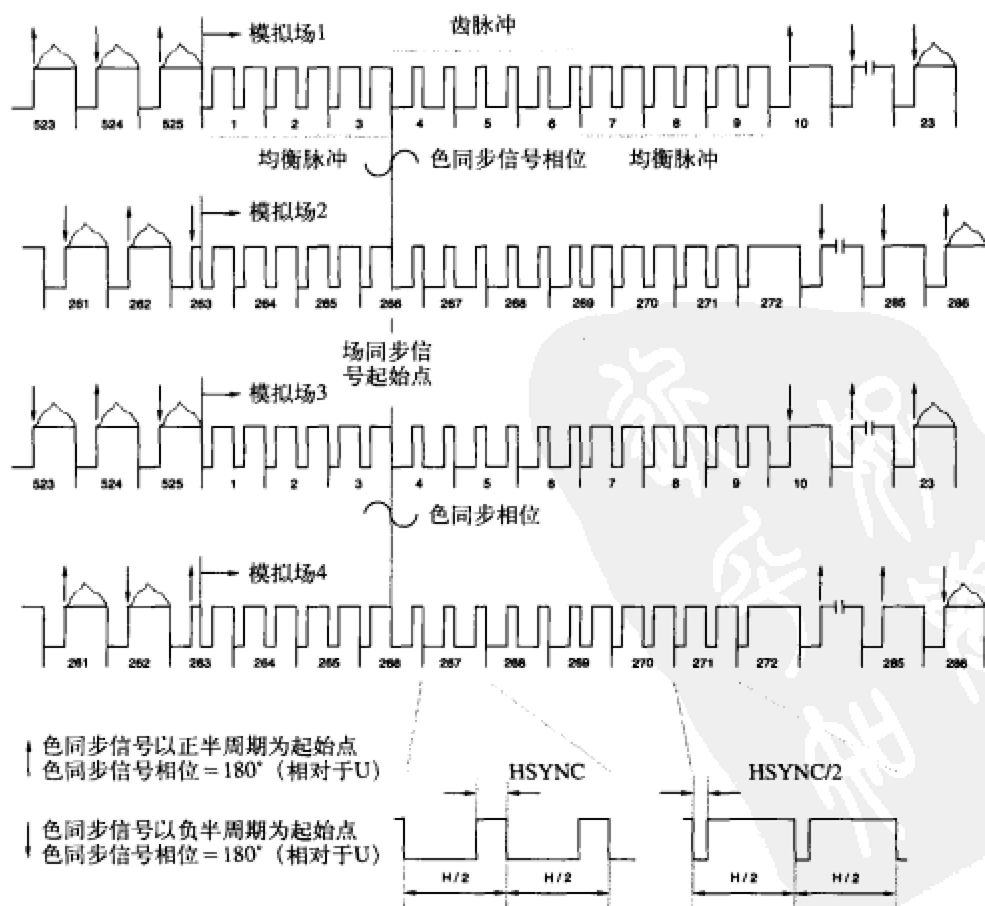


图8-5 4场（M）NTSC序列和色同步信号消隐

8.1.6 NTSC标准

图8-6示范了几种常用的NTSC制式表示方法。字母M代表水平频和场频分别为525和59.94的黑白电视标准，它的视频带宽为4.2MHz，音频载波频率位于视频载波频率以上4.5MHz处，RF频道带宽为6MHz。NTSC指的是将彩色信息加入到黑白信号中的一种技术。表8-9中提供了详细的时序参数。

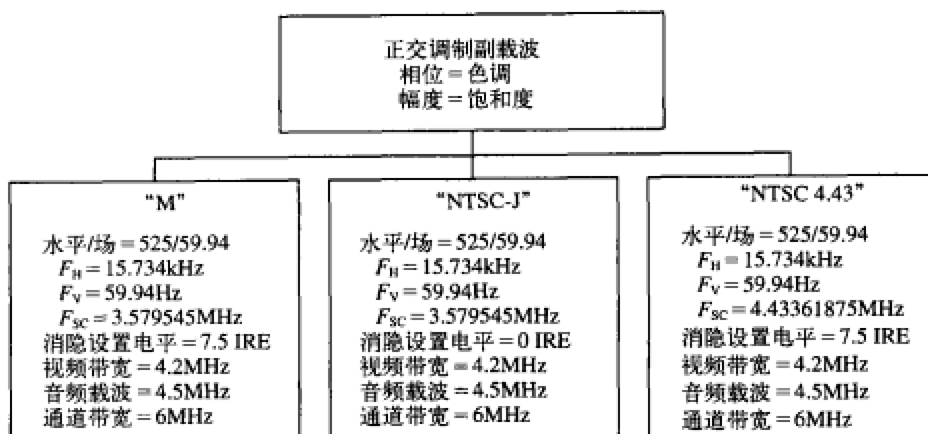


图8-6 常见的NTSC制式

261 ~ 264

多标准的模拟VCR通常使用NTSC 4.43。水平和场的时序和（M）NTSC一致，彩色编码使用PAL调制格式，彩色副载波频率为4.43361875MHz。

NTSC-J用于日本，除了在有效视频区间内没有消隐基底电平以外，其他和（M）NTSC一致。因此，有效视频的标称幅度值为714mV。

逐行NTSC是一种262行、帧频为60帧/秒的NTSC，如图8-7所示。除了每帧有262行以外，这种格式和标准的（M）NTSC完全一致。

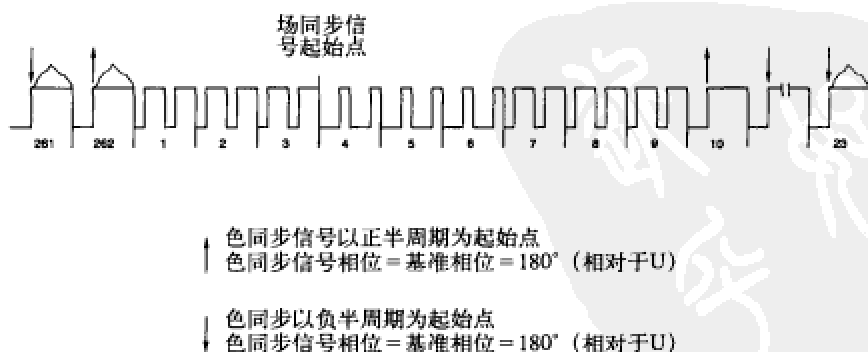
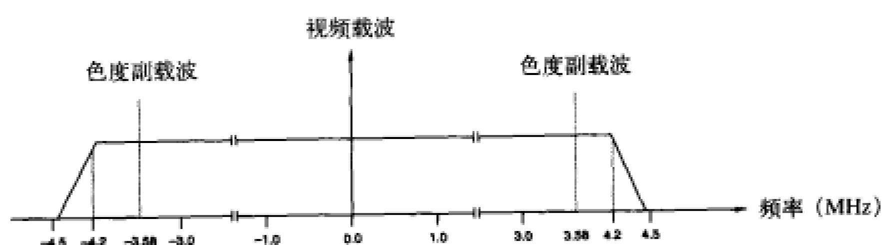


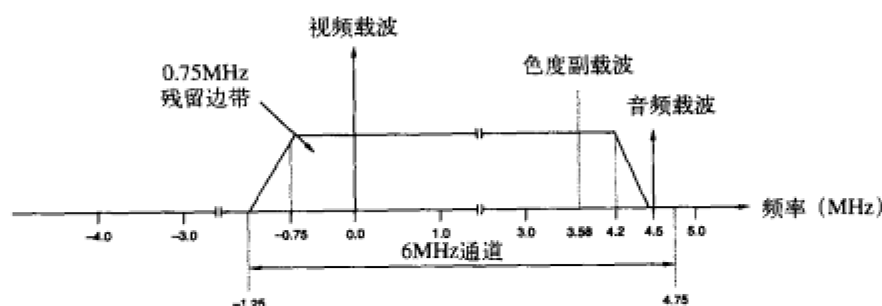
图8-7 逐行NTSC帧序

8.1.7 RF调制

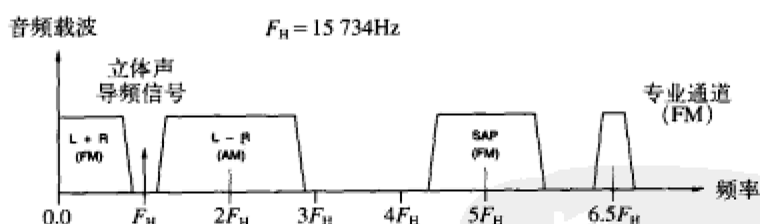
图8-8、图8-9和图8-10描述了基带（M）NTSC复合视频转化为RF（Radio Frequency，射频）信号的基本处理过程。



（a）基带复合视频信号的频谱



（b）包含音频信息的典型通道的频谱



（c）BTSC立体声音频信息的详细频谱

图8-8 （M）NTSC制式的传输通道

图8-8a示范了基带复合视频信号的频谱，它和图8-3类似。然而，图8-3出于简单考虑，只示范了上边带。0MHz位置处的“视频载波”记号只是作为一个与图8-8 b进行比较的参考点。

图8-8b示范了位于一个6MHz通道（例如3频道）内的音频/视频信号。视频信号通过低通滤波，滤去大部分下边带，再加入音频信息。

图8-8c详细描述了针对立体声（BTSC）运算的音频副载波携带的信息。

如图8-9和图8-10所示，模拟视频信号的后肩钳位（back porch clamping）（参见术语表）保证了后肩电平为常数，而与平均图像电平的变化无关。视频信号的白峰限幅（white clipping）防止调制信号低于10%，低于10%可能导致过调制，在电视接收机上产生蜂音。视频信号然后通过4.2MHz的低通滤波器滤波，随后驱动AM（amplitude modulation，幅度调制）视频调制器。同步电平对应于100%调制，消隐电平对应于75%调制，而白色电平对应于10%调制。（M）NTSC制式为视频使用的IF（intermediate frequency，中频）为45.75MHz。

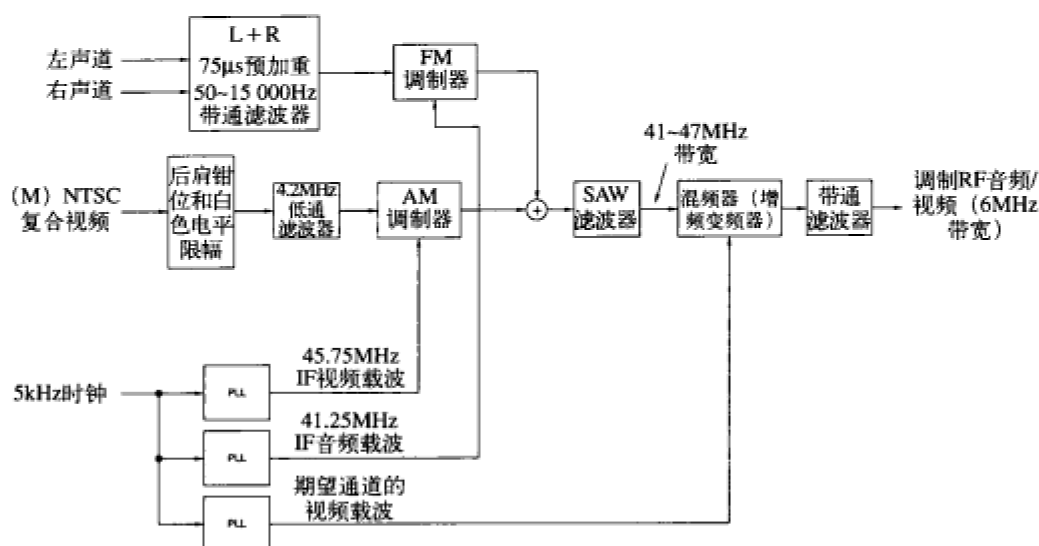


图8-9 (M) NTSC的典型RF调制实现：单声道音频

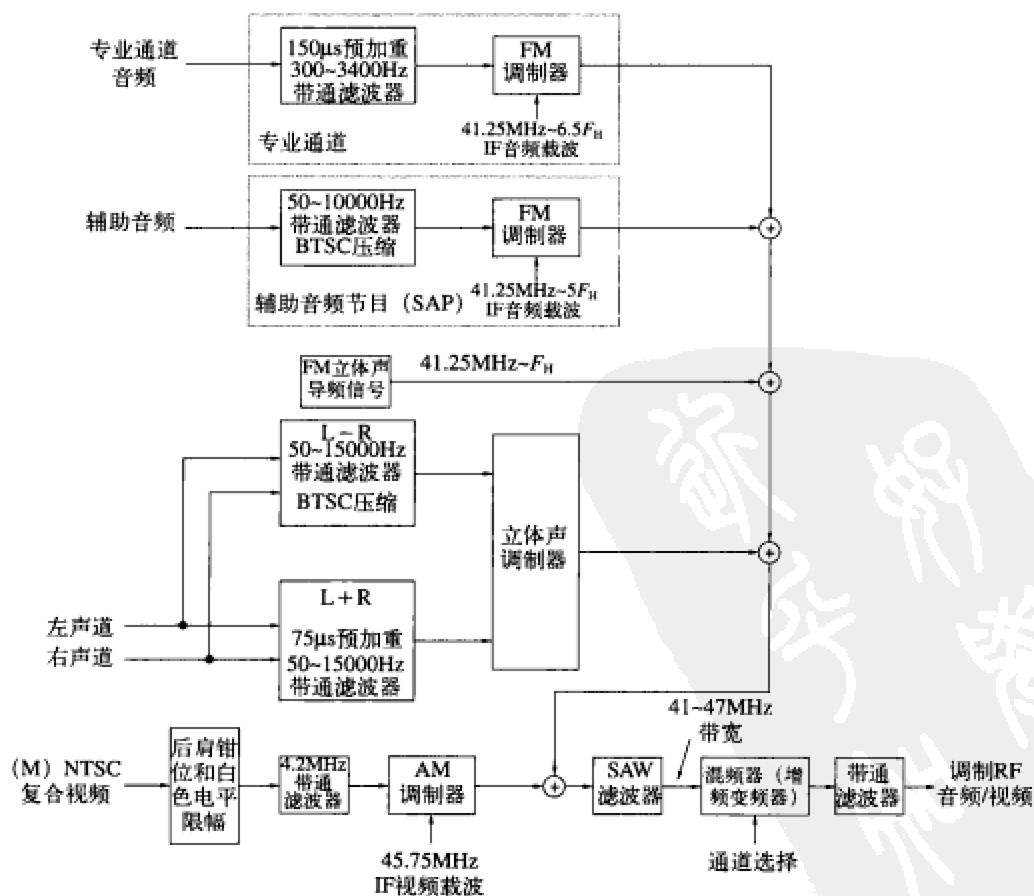


图8-10 (M) NTSC的典型RF调制实现：BTSC立体声音频

在这点上，音频信息叠加在副载波41.25MHz处。如图8-9所示，单声道音频信号被处理后去

驱动FM（Frequency Modulation，频率调制）调制器，然后将FM调制器的输出叠加到IF视频信号中。

作为残留边带滤波器，SAW滤波器滤波IF信号。混频器（或者增频变频器）将IF信号和期望的广播频率混合起来。混频过程中将产生和频与差频，因此使用一个带通滤波器将差频信号抽取出来。

立体声音频（模拟）

❑ BTSC

该标准由EIA TVSB5定义，被公认为BTSC（Broadcast Television Systems Committee，广播电视制式委员会）制式标准，如图8-10所示。使用这种制式的国家包括美国、加拿大、墨西哥和巴西等。

为了使能立体声系统，L－R信息使用一个受抑制的AM副载波传送。有可能存在一个SAP（Secondary Audio Program，辅助音频节目）通道，用来传送一种辅助语言或视频描述（为视觉受损者提供的描述性音频）。也可能存在一个专业通道，允许与远程设备和人进行交流。

❑ Zweiton M

这个标准（ITU-R BS.707），通常也叫作A2 M，和PAL制式中使用的标准类似。L＋R信息通过FM传送，副载波频率为4.5MHz。L－R信息，或者辅助L＋R音频信号，也是通过FM传送的，但它的副载波为4.724 212MHz。

若立体声信号或双路单声道信号存在，那么频率为4.724 212MHz的FM副载波要再经过55.069 9kHz副载波的幅度调制。这一55.069 9kHz副载波或者在149.9Hz处使用50%幅度调制表示立体声音频，或者在276.0Hz处使用50%幅度调制表示双路单声道音频。

这一制式在韩国使用。

266

❑ EIA-J

这一标准与BTSC相似，日本使用这种标准。L＋R信息通过FM传送，副载波频率为4.5MHz。L－R信号，或者辅助L＋R音频信号，也是通过FM传送，但它的副载波为$+2F_H$。

若立体声信号或双路单声道信号存在，那么$+3.5F_H$的副载波或者使用982.5Hz的副载波幅度调制（立体声音频），或者使用922.5Hz的副载波幅度调制（双路单声道音频）。

8.1.8 模拟频道分配

表8-1到表8-4列举了各种NTSC制式的VHF、UHF和有线电视的典型频道分配情况。

值得一提的是，有线电视系统通常会重新分配频道号来改变频率以将干扰最小化，并且提供不同级别的节目内容（例如一般节目和预付费电影频道）。

267～269

表8-1 北美模拟广播标称频率

广播频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)	广播频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)
–	–	–	–	7	175.25	179.75	174~180
–	–	–	–	8	181.25	185.75	180~186
2	55.25	59.75	54~60	9	187.25	191.75	186~192
3	61.25	65.75	60~66	10	193.25	197.75	192~198
4	67.25	71.75	66~72	11	199.25	203.75	198~204
5	77.25	81.75	76~82	12	205.25	209.75	204~210
6	83.25	87.75	82~88	13	211.25	215.75	210~216

（续）

广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
14	471.25	475.75	470~476	42	639.25	643.75	638~644
15	477.25	481.75	476~482	43	645.25	649.75	644~650
16	483.25	487.75	482~488	44	651.25	655.75	650~656
17	489.25	493.75	488~494	45	657.25	661.75	656~662
18	495.25	499.75	494~500	46	663.25	667.75	662~668
19	501.25	505.75	500~506	47	669.25	673.75	668~674
20	507.25	511.75	506~512	48	675.25	679.75	674~680
21	513.25	517.75	512~518	49	681.25	685.75	680~686
22	519.25	523.75	518~524	50	687.25	691.75	686~692
23	525.25	529.75	524~530	51	693.25	697.75	692~698
24	531.25	535.75	530~536	52	699.25	703.75	698~704
25	537.25	541.75	536~542	53	705.25	709.75	704~710
26	543.25	547.75	542~548	54	711.25	715.75	710~716
27	549.25	553.75	548~554	55	717.25	721.75	716~722
28	555.25	559.75	554~560	56	723.25	727.75	722~728
29	561.25	565.75	560~566	57	729.25	733.75	728~734
30	567.25	571.75	566~572	58	735.25	739.75	734~740
31	573.25	577.75	572~578	59	741.25	745.75	740~746
32	579.25	583.75	578~584	60	747.25	751.75	746~752
33	585.25	589.75	584~590	61	753.25	757.75	752~758
34	591.25	595.75	590~596	62	759.25	763.75	758~764
35	597.25	601.75	596~602	63	765.25	769.75	764~770
36	603.25	607.75	602~608	64	771.25	775.75	770~776
37	609.25	613.75	608~614	65	777.25	781.75	776~782
38	615.25	619.75	614~620	66	783.25	787.75	782~788
39	621.25	625.75	620~626	67	789.25	793.75	788~794
40	627.25	631.75	626~632	68	795.25	799.75	794~800
41	633.25	637.75	632~638	69	801.25	805.75	800~806

270

表8-2 日本模拟广播标称频率

广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
–	–	–	–	10	205.25	209.75	204~210
1	91.25	95.75	90~96	11	211.25	215.75	210~216
2	97.25	101.75	96~102	12	217.25	221.75	216~222
3	103.25	107.75	102~108	13	471.25	475.75	470~476
4	171.25	175.75	170~176	14	477.25	481.75	476~482
5	177.25	181.75	176~182	15	483.25	487.75	482~488
6	183.25	187.75	182~188	16	489.25	493.75	488~494
7	189.25	193.75	188~194	17	495.25	499.75	494~500
8	193.25	197.75	192~198	18	501.25	505.75	500~506
9	199.25	203.75	198~204	19	507.25	511.75	506~512

（续）

广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	广播频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
20	513.25	517.75	512~518	45	663.25	667.75	662~668
21	519.25	523.75	518~524	46	669.25	673.75	668~674
22	525.25	529.75	524~530	47	675.25	679.75	674~680
23	531.25	535.75	530~536	48	681.25	685.75	680~686
24	537.25	541.75	536~542	49	687.25	691.75	686~692
25	543.25	547.75	542~548	50	693.25	697.75	692~698
26	549.25	553.75	548~554	51	699.25	703.75	698~704
27	555.25	559.75	554~560	52	705.25	709.75	704~710
28	561.25	565.75	560~566	53	711.25	715.75	710~716
29	567.25	571.75	566~572	54	717.25	721.75	716~722
30	573.25	577.75	572~578	55	723.25	727.75	722~728
31	579.25	583.75	578~584	56	729.25	733.75	728~734
32	585.25	589.75	584~590	57	735.25	739.75	734~740
33	591.25	595.75	590~596	58	741.25	745.75	740~746
34	597.25	601.75	596~602	59	747.25	751.75	746~752
35	603.25	607.75	602~608	60	753.25	757.75	752~758
36	609.25	613.75	608~614	61	759.25	763.75	758-764
37	615.25	619.75	614~620	62	765.25	769.75	764~770
38	621.25	625.75	620~626	—	—	—	—
39	627.25	631.75	626~632	—	—	—	—
40	633.25	637.75	632~638	—	—	—	—
41	639.25	643.75	638~644	—	—	—	—
42	645.25	649.75	644~650	—	—	—	—
43	651.25	655.75	650~656	—	—	—	—
44	657.25	661.75	656~662	—	—	—	—

271

表8-3a 美国标准模拟有线电视标称频率

有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
—	—	—	—	13	211.25	215.75	210~216
—	—	—	—	14	121.2625	125.7625	120~126
2	55.25	59.75	54~60	15	127.2625	131.7625	126~132
3	61.25	65.75	60~66	16	133.2625	137.7625	132~138
4	67.25	71.75	66~72	17	139.25	143.75	138~144
5	77.25	81.75	76~82	18	145.25	149.75	144~150
6	83.25	87.75	82~88	19	151.25	155.75	150~156
7	175.25	179.75	174~180	20	157.25	161.75	156~162
8	181.25	185.75	180~186	21	163.25	167.75	162~168
9	187.25	191.75	186~192	22	169.25	173.75	168~174
10	193.25	197.75	192~198	23	217.25	221.75	216~222
11	199.25	203.75	198~204	24	223.25	227.75	222~228
12	205.25	209.75	204~210	25	229.2625	233.7625	228~234

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
26	235.2625	239.7625	234~240	68	487.25	491.75	486~492
27	241.2625	245.7625	240~246	69	493.25	497.75	492~498
28	247.2625	251.7625	246~252	70	499.25	503.75	498~504
29	253.2625	257.7625	252~258	71	505.25	509.75	504~510
30	259.2625	263.7625	258~264	72	511.25	515.75	510~516
31	265.2625	269.7625	264~270	73	517.25	521.75	516~522
32	271.2625	275.7625	270~276	74	523.25	527.75	522~528
33	277.2625	281.7625	276~282	75	529.25	533.75	528~534
34	283.2625	287.7625	282~288	76	535.25	539.75	534~540
35	289.2625	293.7625	288~294	77	541.25	545.75	540~546
36	295.2625	299.7625	294~300	78	547.25	551.75	546~552
37	301.2625	305.7625	300~306	79	553.25	557.75	552~558
38	307.2625	311.7625	306~312	80	559.25	563.75	558~564
39	313.2625	317.7625	312~318	81	565.25	569.75	564~570
40	319.2625	323.7625	318~324	82	571.25	575.75	570~576
41	325.2625	329.7625	324~330	83	577.25	581.75	576~582
42	331.2750	335.7750	330~336	84	583.25	587.75	582~588
43	337.2625	341.7625	336~342	85	589.25	593.75	588~594
44	343.2625	347.7625	342~348	86	595.25	599.75	594~600
45	349.2625	353.7625	348~354	87	601.25	605.75	600~606
46	355.2625	359.7625	354~360	88	607.25	611.75	606~612
47	361.2625	365.7625	360~366	89	613.25	617.75	612~618
48	367.2625	371.7625	366~372	90	619.25	623.75	618~624
49	373.2625	377.7625	372~378	91	625.25	629.75	624~630
50	379.2625	383.7625	378~384	92	631.25	635.75	630~636
51	385.2625	389.7625	384~390	93	637.25	641.75	636~642
52	391.2625	395.7625	390~396	94	643.25	647.75	642~648
53	397.2625	401.7625	396~402	95	91.25	95.75	90~96
54	403.25	407.75	402~408	96	97.25	101.75	96~102
55	409.25	413.75	408~414	97	103.25	107.75	102~108
56	415.25	419.75	414~420	98	109.2750	113.7750	108~114
57	421.25	425.75	420~426	99	115.2750	119.7750	114~120
58	427.25	431.75	426~432	100	649.25	653.75	648~654
59	433.25	437.75	432~438	101	655.25	659.75	654~660
60	439.25	443.75	438~444	102	661.25	665.75	660~666
61	445.55	449.75	444~450	103	667.25	671.75	666~672
62	451.25	455.75	450~456	104	673.25	677.75	672~678
63	457.25	461.75	456~462	105	679.25	683.75	678~684
64	463.25	467.75	462~468	106	685.25	689.75	684~690
65	469.25	473.75	468~474	107	691.25	695.75	690~696
66	475.25	479.75	474~480	108	697.25	701.75	696~702
67	481.25	485.75	480~486	109	703.25	707.75	702~708

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
110	709.25	713.75	708~714	135	859.25	863.75	858~864
111	715.25	719.75	714~720	136	865.25	869.75	864~870
112	721.25	725.75	720~726	137	871.25	875.75	870~876
113	727.25	731.75	726~732	138	877.25	881.75	876~882
114	733.25	737.75	732~738	139	883.25	887.75	882~888
115	739.25	743.75	738~744	140	889.25	893.75	888~894
116	745.25	749.75	744~750	141	895.25	899.75	894~900
117	751.25	755.75	750~756	142	901.25	905.75	900~906
118	757.25	761.75	756~762	143	907.25	911.75	906~912
119	763.25	767.75	762~768	144	913.25	917.75	912~918
120	769.25	773.75	768~774	145	919.25	923.75	918~924
121	775.25	779.75	774~780	146	925.25	929.75	924~930
122	781.25	785.75	780~786	147	931.25	935.75	930~936
123	787.25	791.75	786~792	148	937.25	941.75	936~942
124	793.25	797.75	792~798	149	943.25	947.75	942-948
125	799.25	803.75	798~804	150	949.25	953.75	948~954
126	805.25	809.75	804~810	151	955.25	959.75	954~960
127	811.25	815.75	810~816	152	961.25	965.75	960~966
128	817.25	821.75	816~822	153	967.25	971.75	966~972
129	823.25	827.75	822~828	154	973.25	977.75	972~978
130	829.25	833.75	828~834	155	979.25	983.75	978~984
131	835.25	839.75	834~840	156	985.25	989.75	984~990
132	841.25	845.75	840~846	157	991.25	995.75	990~996
133	847.25	851.75	846~852	158	997.25	1001.75	996~1002
134	853.25	857.75	852~858	–	–	–	–

表8-3b 美国模拟有线电视标称频率：增量相关载波（Incrementally Related Carrier，IRC）制式

有线频道	视频载波（MHz）	音频载波（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）
–	–	–	13	211.2625	215.7625
1	73.2625	77.7625	14	121.2625	125.7625
2	55.2625	59.7625	15	127.2625	131.7625
3	61.2625	65.7625	16	133.2625	137.7625
4	67.2625	71.7625	17	139.2625	143.7625
5	79.2625	83.7625	18	145.2625	149.7625
6	85.2625	89.7625	19	151.2625	155.7625
7	175.2625	179.7625	20	157.2625	161.7625
8	181.2625	185.7625	21	163.2625	167.7625
9	187.2625	191.7625	22	169.2625	173.7625
10	193.2625	197.7625	23	217.2625	221.7625
11	199.2625	203.7625	24	223.2625	227.7625
12	205.2625	209.7625	25	229.2625	233.7625

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）
26	235.2625	239.7625	68	487.2625	491.7625
27	241.2625	245.7625	69	493.2625	497.7625
28	247.2625	251.7625	70	499.2625	503.7625
29	253.2625	257.7625	71	505.2625	509.7625
30	259.2625	263.7625	72	511.2625	515.7625
31	265.2625	269.7625	73	517.2625	521.7625
32	271.2625	275.7625	74	523.2625	527.7625
33	277.2625	281.7625	75	529.2625	533.7625
34	283.2625	287.7625	76	535.2625	539.7625
35	289.2625	293.7625	77	541.2625	545.7625
36	295.2625	299.7625	78	547.2625	551.7625
37	301.2625	305.7625	79	553.2625	557.7625
38	307.2625	311.7625	80	559.2625	563.7625
39	313.2625	317.7625	81	565.2625	569.7625
40	319.2625	323.7625	82	571.2625	575.7625
41	325.2625	329.7625	83	577.2625	581.7625
42	331.2750	335.7750	84	583.2625	587.7625
43	337.2625	341.7625	85	589.2625	593.7625
44	343.2625	347.7625	86	595.2625	599.7625
45	349.2625	353.7625	87	601.2625	605.7625
46	355.2625	359.7625	88	607.2625	611.7625
47	361.2625	365.7625	89	613.2625	617.7625
48	367.2625	371.7625	90	619.2625	623.7625
49	373.2625	377.7625	91	625.2625	629.7625
50	379.2625	383.7625	92	631.2625	635.7625
51	385.2625	389.7625	93	637.2625	641.7625
52	391.2625	395.7625	94	643.2625	647.7625
53	397.2625	401.7625	95	91.2625	95.7625
54	403.2625	407.7625	96	97.2625	101.7625
55	409.2625	413.7625	97	103.2625	107.7625
56	415.2625	419.7625	98	109.2750	113.7750
57	421.2625	425.7625	99	115.2625	119.7625
58	427.2625	431.7625	100	649.2625	653.7625
59	433.2625	437.7625	101	655.2625	659.7625
60	439.2625	443.7625	102	661.2625	665.7625
61	445.2625	449.7625	103	667.2625	671.7625
62	451.2625	455.7625	104	673.2625	677.7625
63	457.2625	461.7625	105	679.2625	683.7625
64	463.2625	467.7625	106	685.2625	689.7625
65	469.2625	473.7625	107	691.2625	695.7625
66	475.2625	479.7625	108	697.2625	701.7625
67	481.2625	485.7625	109	703.2625	707.7625

272

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）
110	709.2625	713.7625	135	859.2625	863.7625
111	715.2625	719.7625	136	865.2625	869.7625
112	721.2625	725.7625	137	871.2625	875.7625
113	727.2625	731.7625	138	877.2625	881.7625
114	733.2625	737.7625	139	883.2625	887.7625
115	739.2625	743.7625	140	889.2625	893.7625
116	745.2625	749.7625	141	895.2625	899.7625
117	751.2625	755.7625	142	901.2625	905.7625
118	757.2625	761.7625	143	907.2625	911.7625
119	763.2625	767.7625	144	913.2625	917.7625
120	769.2625	773.7625	145	919.2625	923.7625
121	775.2625	779.7625	146	925.2625	929.7625
122	781.2625	785.7625	147	931.2625	935.7625
123	787.2625	791.7625	148	937.2625	941.7625
124	793.2625	797.7625	149	943.2625	947.7625
125	799.2625	803.7625	150	949.2625	953.7625
126	805.2625	809.7625	151	955.2625	959.7625
127	811.2625	815.7625	152	961.2625	965.7625
128	817.2625	821.7625	153	967.2625	971.7625
129	823.2625	827.7625	154	973.2625	977.7625
130	829.2625	833.7625	155	979.2625	983.7625
131	835.2625	839.7625	156	985.2625	989.7625
132	841.2625	845.7625	157	991.2625	995.7625
133	847.2625	851.7625	158	997.2625	1001.7625
134	853.2625	857.7625	–	–	–

表8-3c 美国模拟有线电视标称频率：和声相关载波（Harmonically Related Carrier，HRC）制式

有线频道	视频载波（MHz）	音频载波（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）
–	–	–	13	210.0105	214.5105
1	72.0036	76.5036	14	120.0060	124.5060
2	54.0027	58.5027	15	126.0063	130.5063
3	60.0030	64.5030	16	132.0066	136.5066
4	66.0033	70.5030	17	138.0069	142.5069
5	72.0036	82.5039	18	144.0072	148.5072
6	78.0039	88.5042	19	150.0075	154.5075
7	174.0087	178.5087	20	156.0078	160.5078
8	180.0090	184.5090	21	162.0081	166.5081
9	186.0093	190.5093	22	168.0084	172.5084
10	192.0096	196.5096	23	216.0108	220.5108
11	198.0099	202.5099	24	222.0111	226.5111
12	204.0102	208.5102	25	228.0114	232.5114

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）
26	234.0117	238.5117	68	486.0243	490.5243
27	240.0120	244.5120	69	492.0246	496.5246
28	246.0123	250.5123	70	498.0249	502.5249
29	252.0126	256.5126	71	504.0252	508.5252
30	258.0129	262.5129	72	510.0255	514.5255
31	264.0132	268.5132	73	516.0258	520.5258
32	270.0135	274.5135	74	522.0261	526.5261
33	276.0138	280.5138	75	528.0264	532.5264
34	282.0141	286.5141	76	534.0267	538.5267
35	288.0144	292.5144	77	540.0270	544.5270
36	294.0147	298.5147	78	546.0273	550.5273
37	300.0150	304.5150	79	552.0276	556.5276
38	306.0153	310.5153	80	558.0279	562.5279
39	312.0156	316.5156	81	564.0282	568.5282
40	318.0159	322.5159	82	570.0285	574.5285
41	324.0162	328.5162	83	576.0288	580.5288
42	330.0165	334.5165	84	582.0291	586.5291
43	336.0168	340.5168	85	588.0294	592.5294
44	342.0168	346.5168	86	594.0297	598.5297
45	348.0168	352.5168	87	600.0300	604.5300
46	354.0168	358.5168	88	606.0303	610.5303
47	360.0168	364.5168	89	612.0306	616.5306
48	366.0168	370.5168	90	618.0309	622.5309
49	372.0168	376.5168	91	624.0312	628.5312
50	378.0168	382.5168	92	630.0315	634.5315
51	384.0168	388.5168	93	636.0318	640.5318
52	390.0168	394.5168	94	642.0321	646.5321
53	396.0168	400.5168	95	90.0045	94.5045
54	402.0201	406.5201	96	96.0048	100.5048
55	408.0204	412.5204	97	102.0051	106.5051
56	414.0207	418.5207	98	–	–
57	420.0210	424.5210	99	–	–
58	426.0213	430.5213	100	648.0324	652.5324
59	432.0216	436.5216	101	654.0327	658.5327
60	438.0219	442.5219	102	660.0330	664.5330
61	444.0222	448.5222	103	666.0333	670.5333
62	450.0225	454.5225	104	672.0336	676.5336
63	456.0228	460.5228	105	678.0339	682.5339
64	462.0231	466.5231	106	684.0342	688.5342
65	468.0234	472.5234	107	690.0345	694.5345
66	474.0237	478.5237	108	696.0348	700.5348
67	480.0240	484.5240	109	702.0351	706.5351

（续）

有线频道	视频载波 (MHz)	音频载波 (MHz)	有线频道	视频载波 (MHz)	音频载波 (MHz)
110	708.0354	712.5354	135	858.0429	862.5429
111	714.0357	718.5357	136	864.0432	868.5432
112	720.0360	724.5360	137	870.0435	874.5435
113	726.0363	730.5363	138	876.0438	880.5438
114	732.0366	736.5366	139	882.0441	888.5441
115	738.0369	742.5369	140	888.0444	892.5444
116	744.0372	748.5372	141	894.0447	898.5447
117	750.0375	754.5375	142	900.0450	904.5450
118	756.0378	760.5378	143	906.0453	910.5453
119	762.0381	766.5381	144	912.0456	916.5456
120	768.0384	772.5384	145	918.0459	922.5459
121	774.0387	778.5387	146	924.0462	928.5462
122	780.0390	784.5390	147	930.0465	934.5465
123	786.0393	790.5393	148	936.0468	940.5468
124	792.0396	796.5396	149	942.0471	946.5471
125	798.0399	802.5399	150	948.0474	952.5474
126	804.0402	808.5402	151	954.0477	958.5477
127	810.0405	814.5405	152	960.0480	964.5480
128	816.0408	820.5408	153	966.0483	970.5483
129	822.0411	826.5411	154	972.0486	976.5486
130	828.0414	832.5414	155	978.0489	982.5489
131	834.0417	838.5417	156	984.0492	988.5492
132	840.0420	844.5420	157	990.0495	994.5495
133	846.0423	850.5423	158	996.0498	1000.5498
134	852.0426	856.5426	–	–	–

274 ~ 277

表8-4 日本模拟有线电视标称频率

有线频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)	有线频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)
–	–	–	–	23	223.25	227.75	222~228
–	–	–	–	24	231.25	235.75	230~236
–	–	–	–	25	237.25	241.75	236~242
13	109.25	113.75	108~114	26	243.25	247.75	242~248
14	115.25	119.75	114~120	27	249.25	253.75	248~254
15	121.25	125.75	120~126	28	253.25	257.75	252~258
16	127.25	131.75	126~132	29	259.25	263.75	258~264
17	133.25	137.75	132~138	30	265.25	269.75	264~270
18	139.25	143.75	138~144	31	271.25	275.75	270~276
19	145.25	149.75	144~150	32	277.25	281.75	276~282
20	151.25	155.75	150~156	33	283.25	287.75	282~288
21	157.25	161.75	156~162	34	289.25	293.75	288~294
22	165.25	169.75	164~170	35	295.25	299.75	294~300

（续）

有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
36	301.25	305.75	300~306	53	403.25	407.75	402~408
37	307.25	311.75	306~312	54	409.25	413.75	408~414
38	313.25	317.75	312~318	55	415.25	419.75	414~420
39	319.25	323.75	318~324	56	421.25	425.75	420~426
40	325.25	329.75	324~330	57	427.25	431.75	426~432
41	331.25	335.75	330~336	58	433.25	437.75	432~438
42	337.25	341.75	336~342	59	439.25	443.75	438~444
43	343.25	347.75	342~348	60	445.25	449.75	444~450
44	349.25	353.75	348~354	61	451.25	455.75	450~456
45	355.25	359.75	354~360	62	457.25	461.75	456~462
46	361.25	365.75	360~366	63	463.25	467.75	462~468
47	367.25	371.75	366~372	—	—	—	—
48	373.25	377.75	372~378	—	—	—	—
49	379.25	383.75	378~384	—	—	—	—
50	385.25	389.75	384~390	—	—	—	—
51	391.25	395.75	390~396	—	—	—	—
52	397.25	401.75	396~402	—	—	—	—

8.1.9　亮度方程推导

由RGB生成亮度的方程由接收机所使用的三基色的色度和真正的白光源决定。

RGB三基色的色度和基准白光（CIE照明C）由1953 NTSC标准指定：

R：$x_r=0.67$　　$y_r=0.33$　　$z_r=0.00$

G：$x_g=0.21$　　$y_g=0.71$　　$z_g=0.08$

B：$x_b=0.14$　　$y_b=0.08$　　$z_b=0.78$

白色：$x_w=0.3101$　　$y_w=0.3162$　　$z_w=0.3737$

其中x和y由CIE 1931色度坐标指定，而z由$x+y+z=1$计算。

亮度由RGB的加权和计算，权重代表生成基准白光的亮度时RGB三基色的实际贡献。我们可以通过求解下面的方程式得出生成基准白光的RGB的线性组合：

$$\begin{bmatrix} x_r & x_g & x_b \\ y_r & y_g & y_b \\ z_r & z_g & z_b \end{bmatrix}\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix}=\begin{bmatrix} x_w/y_w \\ 1 \\ z_w/y_w \end{bmatrix}$$

对上述方程式重新排列以求解K_r、K_g和K_b：

$$\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix}=\begin{bmatrix} x_w/y_w \\ 1 \\ z_w/y_w \end{bmatrix}\begin{bmatrix} x_r & x_g & x_b \\ y_r & y_g & y_b \\ z_r & z_g & z_b \end{bmatrix}^{-1}$$

将已知数据代入上式，得出K_r、K_g和K_b：

$$\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix} = \begin{bmatrix} 0.3101/0.3162 \\ 1 \\ 0.3737/0.3162 \end{bmatrix} \begin{bmatrix} 0.67 & 0.21 & 0.14 \\ 0.33 & 0.71 & 0.08 \\ 0.00 & 0.08 & 0.78 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 0.9807 \\ 1 \\ 1.1818 \end{bmatrix} \begin{bmatrix} 1.730 & -0.482 & -0.261 \\ -0.814 & 1.652 & -0.023 \\ 0.083 & -0.169 & 1.284 \end{bmatrix}$$

$$= \begin{bmatrix} 0.906 \\ 0.827 \\ 1.430 \end{bmatrix}$$

Y定义为：

$$Y = (K_r y_r)R' + (K_g y_g)G' + (K_b y_b)B'$$
$$= (0.906)(0.33)R' + (0.827)(0.71)G' + (1.430)(0.08)B'$$

或者：

$$Y = 0.299R' + 0.587G' + 0.114B'$$

现代的接收机使用一组不同的RGB磷光粉，从而使产生的RGB三基色的色度和基准白光（CIE照明D_{65}）稍微有所不同：

R：$x_r = 0.630$　　$y_r = 0.340$　　$z_r = 0.030$

G：$x_g = 0.310$　　$y_g = 0.595$　　$z_g = 0.095$

B：$x_b = 0.155$　　$y_b = 0.070$　　$z_b = 0.775$

白色：$x_w = 0.3127$　　$y_w = 0.3290$　　$z_w = 0.3583$

其中x和y由CIE 1931色度坐标指定；而z由$x + y + z = 1$计算。将已知数据代入后，可以得到K_r、K_g和K_b：

279

$$\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix} = \begin{bmatrix} 0.3127/0.3290 \\ 1 \\ 0.3583/0.3290 \end{bmatrix} \begin{bmatrix} 0.630 & 0.310 & 0.155 \\ 0.340 & 0.595 & 0.070 \\ 0.030 & 0.095 & 0.775 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 0.6243 \\ 1.1770 \\ 1.2362 \end{bmatrix}$$

因为Y定义为：

$$Y = (K_r y_r)R' + (K_g y_g)G' + (K_b y_b)B'$$
$$= (0.6243)(0.340)R' + (1.1770)(0.595)G' + (1.2362)(0.070)B'$$

所以结果为：

$$Y = 0.212R' + 0.700G' + 0.086B'$$

但是，仍然可以使用标准的$Y = 0.299R' + 0.587G' + 0.114B'$方程式。为了减少色彩误差，需要对接收机进行调整。

8.2 PAL概述

在彩色电视制式标准上，欧洲没有立即采纳某一种标准，而是评估了1953年至1967年之间与

他们的625行、50场/秒、2：1逐行黑白电视标准兼容的各种制式。欧洲对NTSC规范进行了修改，以克服NTSC在广播过程中的高阶相位和幅度完整性要求，从而避免色彩失真。逐行倒相（Phase Alternation Line，PAL）制式最初依赖于人眼对色彩失真的视觉平均补偿作用，实现了一种逐行倒相的彩色分量。PAL制式于1967年在德国和英国开始广播，这两个国家分别使用了一种稍微有所不同的PAL制式标准。

8.2.1　亮度信息

由R′G′B′推导的黑白亮度（Y）信号为：

$Y = 0.299R' + 0.587G' + 0.114B'$

和NTSC制式一样，PAL制式的亮度信号也占用整个视频带宽。PAL制式有若干种变体，每种变体的视频带宽和音频副载波的位置有所不同。根据不同的PAL制式标准，复合视频信号的带宽有4.2MHz、5.0MHz、5.5MHz和6.0MHz几种。

280

8.2.2　彩色信息

在传送彩色信息时，使用的是U和V：

$U = 0.492(B' - Y)$

$V = 0.877(R' - Y)$

其中U和V的典型带宽为1.3MHz。

8.2.3　彩色调制

和NTSC制式中一样，U和V使用两个相位正交的平衡调制器来调制彩色副载波：其中一个调制器由正弦相位的副载波驱动，另一个调制器由余弦相位的副载波驱动。两个调制器的输出叠加在一起形成色度调制信号：

$C = U\sin\omega t \pm V\cos\omega t$

$\omega = 2\pi F_{SC}$

对于（B、D、G、H、I、N）PAL：$F_{SC} = 4.433\,618\,75\text{MHz}$（±5Hz）

对于（N_C）PAL：$F_{SC} = 3.582\,056\,25\text{MHz}$（±5Hz）

对于（M）PAL：$F_{SC} = 3.575\,611\,49\text{MHz}$（±10Hz）

在PAL制式中，V的相位是逐行倒相的。选择V进行倒相处理是因为它相对于U拥有较低的增益系数，因此它对（1/2）F_H交换频率不平衡较不敏感。V相位以行频交换的结果就是所有的彩色副载波相位误差生成互补的误差，允许在接收机处逐行平均，消除误差并生成饱和度稍微下降的正确色调。这一技术要求PAL接收机能够确定正确的V相位。这一问题可以通过著名的AB同步、PAL同步、PAL开关或者摇摆色同步信号（swinging burst）技术（包括色同步信号的相位以行频±45°的交换）完成。PAL制的UV矢量图如图8-11和图8-12所示。

简单PAL解码器依赖于人眼对逐行色调错误的平均补偿作用。标准PAL解码器使用一个1行延迟线在求均值过程中分离U和V。两种实现都有Hanover条问题，在这种问题中，每对邻接的行都有一个实的和一个互补的色调误差。行求均值过程所造成的结果就是使色度垂直解析度下降。

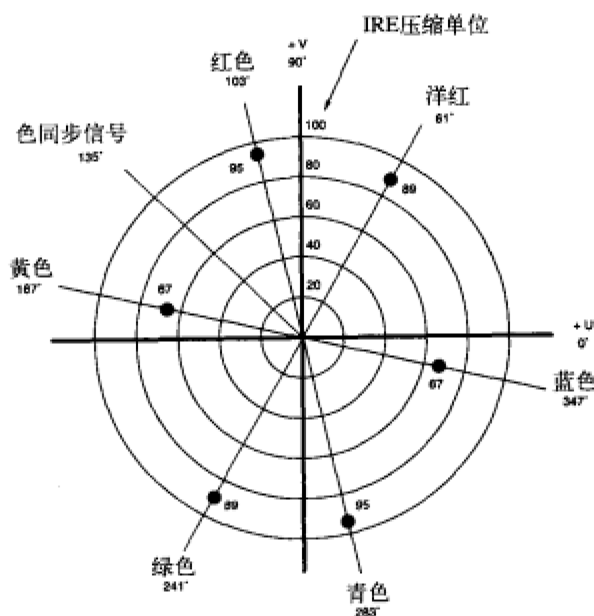


图8-11 75%彩条的UV矢量图。行[n]、PAL开关=0

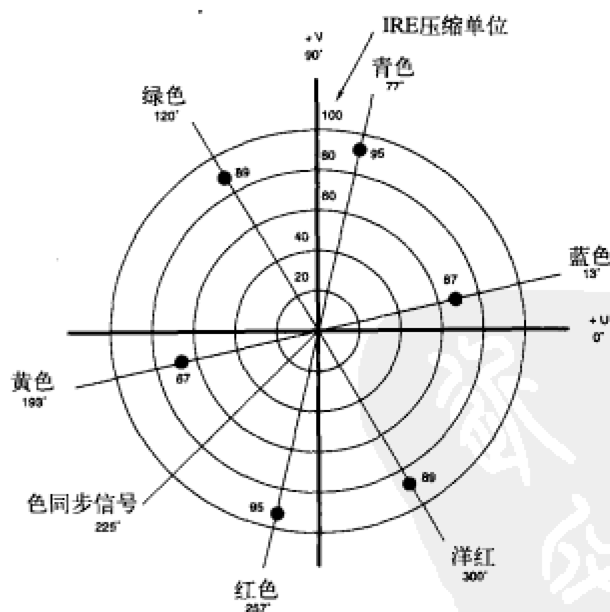


图8-12 75%彩条的UV矢量图。行[n+1]、PAL开关=1

8.2.4 复合视频的生成

调制色度信息加上亮度信息以及合适的水平和场同步信号、消隐信号及色同步信号就生成复合彩色视频波形，如图8-13所示。

复合PAL = Y + Usin ωt ± Vcos ωt + 时序

所生成的复合视频信号的带宽如图8-14所示。

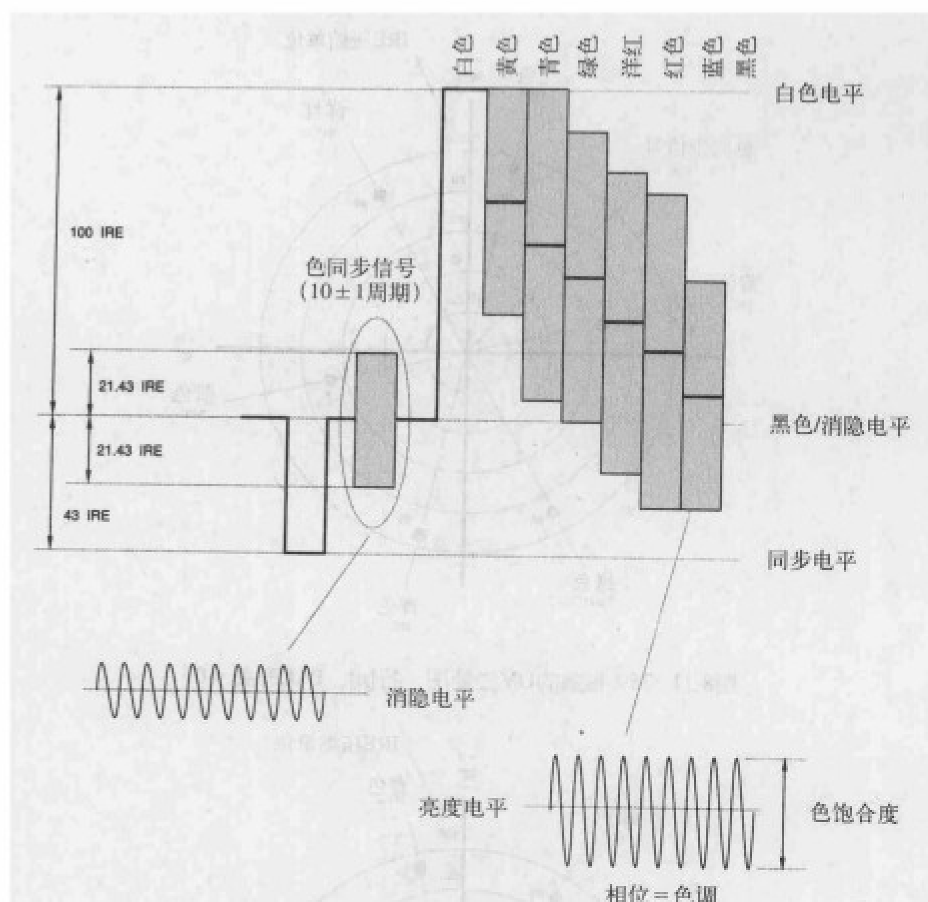


图8-13　75%彩条的（B、D、G、H、I、N_C）PAL复合视频信号



（I）PAL



（B、G、H）PAL

图8-14　一些PAL制式的视频带宽

和NTSC制式相似，由于水平消隐，亮度分量也相隔F_H间隔。因为V分量以1/2行频对称切换，所以只生成奇数的谐频，致使V分量相隔F_H间隔。V分量和U分量错开半行间隔，而U分量也拥有

F_H间隔。如果和NTSC制式一样，副载波使用半行频偏置，那么U分量将完全隔行，但是V分量将和Y分量完全重叠，因此就无法隔行，产生垂直静态点模式。基于这个原因，所以PAL制式副载波频率使用1/4行频偏置：

281

对于（B、D、G、H、I、N）PAL：$F_{SC}=[(1135/4)+(1/625)]\ F_H$

对于（M）PAL：$F_{SC}=(909/4)F_H$

对于（N_C）PAL：$F_{SC}=[(917/4)+(1/625)]\ F_H$

在625行扫描标准下，PAL副载波频率还附加了一个（1/625）F_H（等于25Hz）因子，增加彩色点模式，减小可视度。图8-15描述了所生成的频谱交错图。重复一个具体的样本位置需要8个完整的场，如图8-16和图8-17所示。

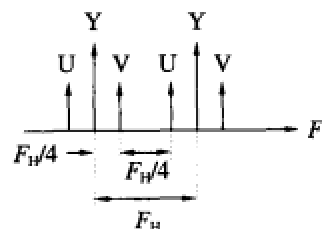


图8-15 亮度和色度频谱交错原理

场同步信号起始点

模拟场1

色同步信号相位的-U分量

模拟场2

模拟场3

模拟场4

色同步信号消隐期

场1

场2

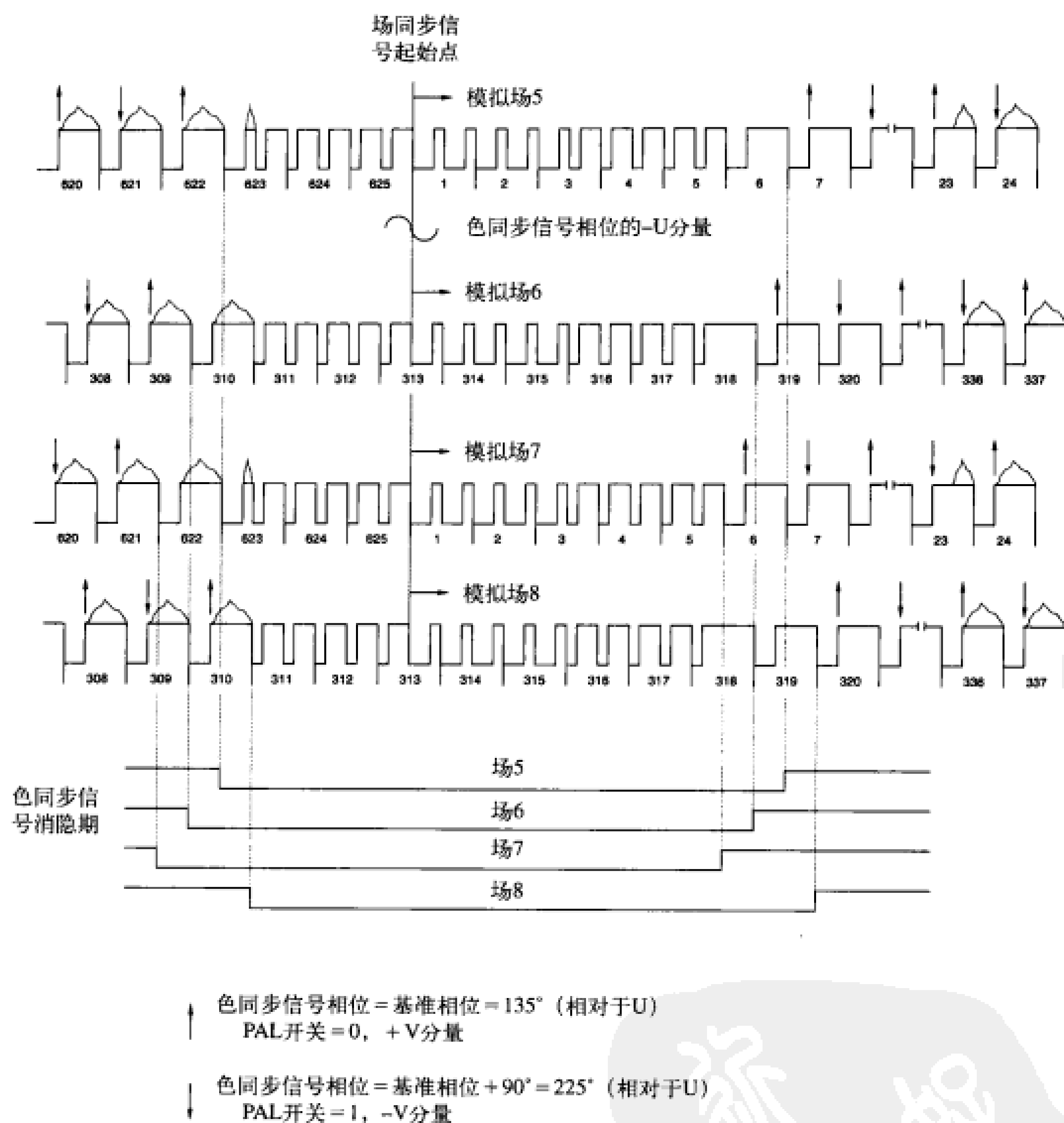
场3

场4

↑ 色同步信号相位＝基准相位＝135°（相对于U）
PAL开关＝0，＋V分量

↓ 色同步信号相位＝基准相位＋90°＝225°（相对于U）
PAL开关＝1，－V分量

图8-16 8场（B、D、G、H、I、N_C）PAL场序和色同步信号消隐。均衡和齿脉冲的细节请参见图8-5



↑ 色同步信号相位＝基准相位＝135°（相对于U）
PAL开关＝0，＋V分量

↓ 色同步信号相位＝基准相位＋90°＝225°（相对于U）
PAL开关＝1，－V分量

图8-16　（续）

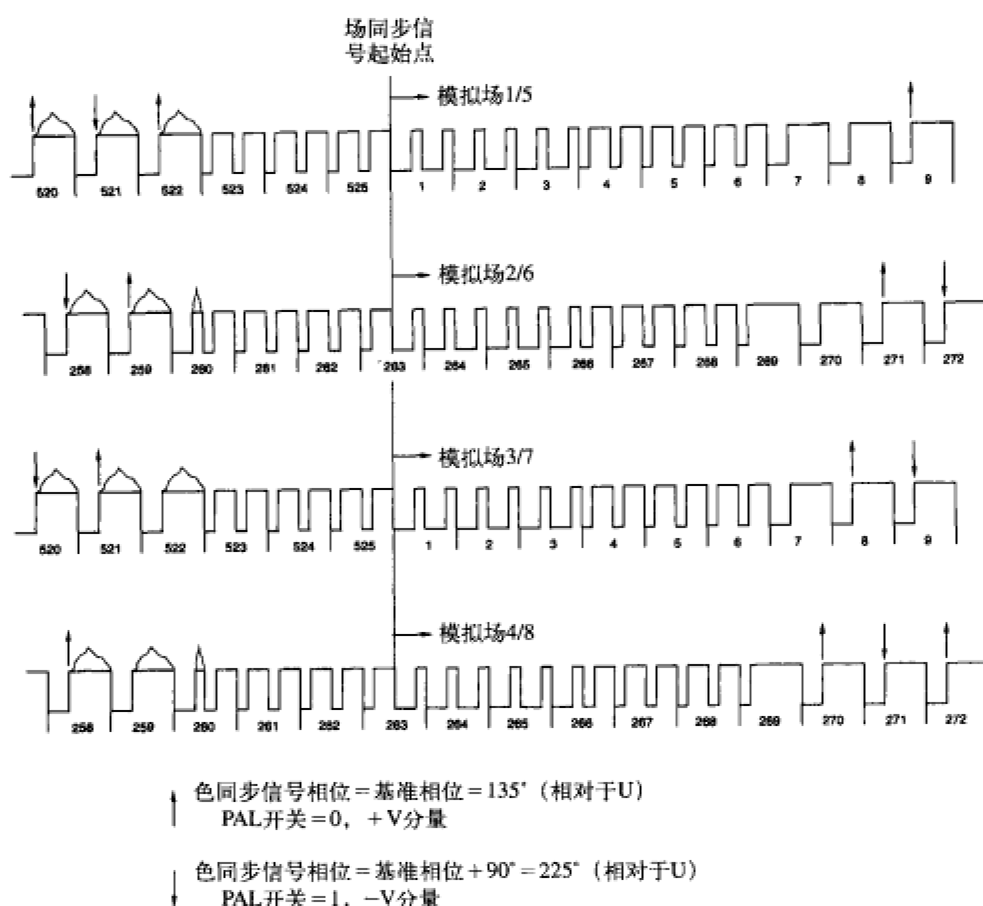


图8-17 8场（M）PAL场序和色同步信号消隐。均衡和齿脉冲的细节请参见图8-5

8.2.5 PAL标准

图8-19示范了几种常用的PAL制式表示方法。各个字母分别代表各种行频场频、视频带宽（4.2MHz、5.0MHz、5.5MHz或6.0MHz）、音频载波相对频率和RF频道带宽（6.0MHz、7.0MHz或8.0MHz）的黑白电视标准。PAL指的是向黑白电视信号添加彩色信息的技术。详细的时序参数位于表8-9中。

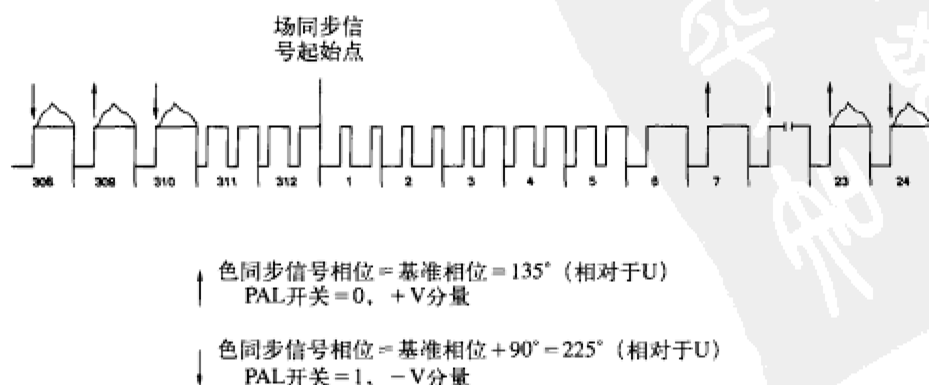


图8-18 逐行PAL帧序

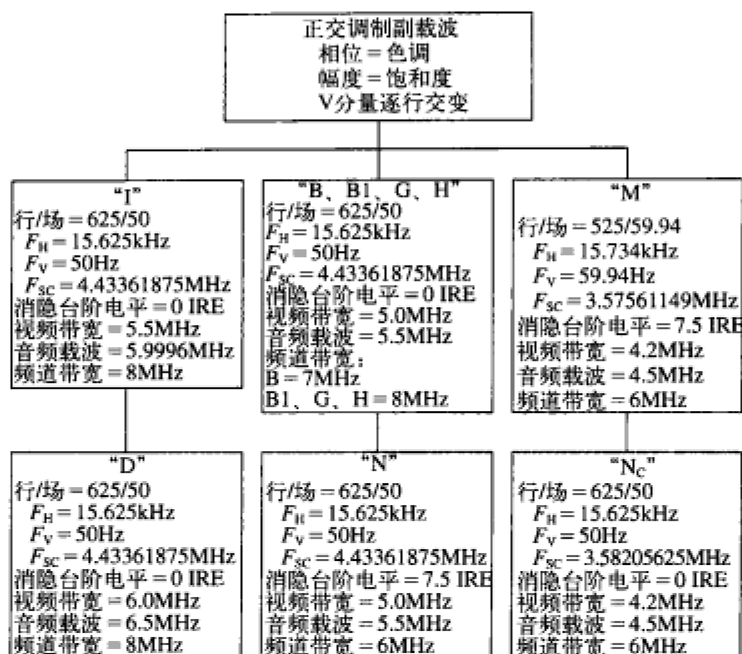


图8-19 常见PAL制式

逐行PAL制式如图8-18所示，是一种312行、50帧/秒的PAL制式，通常应用在视频游戏和银幕显示中。除了每帧只有312行之外，该格式与标准PAL制式相同。

8.2.6 RF调制

图8-20和图8-21描述了基带（G）PAL复合视频向RF（Radio Frequency，射频）信号的转换过程。基本上除了在视频带宽和副载波频率上有所不同外，其他PAL标准的转换过程与该过程相似。

图8-20a示范了一个（G）PAL基带复合视频信号的频谱。它与图8-14相似。但是，图8-14出于简化性，只示范了上边带。0MHz位置处的“视频载波”记号只是作为一个和图8-20 b 进行比较的基准点。

图8-20b示范了位于一个8MHz通道内的音频/视频信号。视频信号通过低通滤波，滤去大多数下边带，再加入音频信息。注意（H）和（I）PAL拥有一个1.25MHz的残留边带，而不是0.75MHz。

图8-20c详细描述了模拟立体声运算的音频副载波携带的信息。

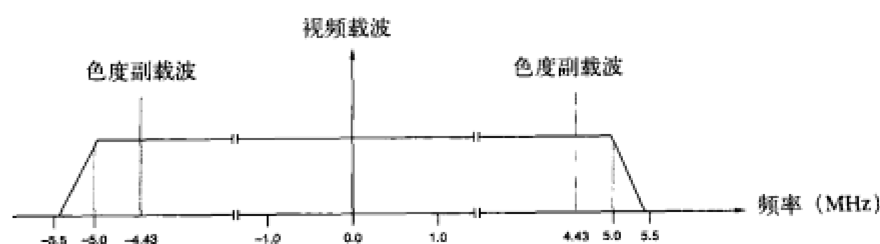
如图8-21所示，模拟视频信号的后肩钳位保证了后肩电平值为常数，而与平均图像电平的变化无关。视频信号然后通过5.0MHz的低通滤波器滤波，随后驱动AM（Amplitude Modulation，幅度调制）视频调制器。同步电平对应于100%调制，根据PAL格式的不同，消隐电平和白色电平对应的调制有所不同：

消隐电平（%调制）

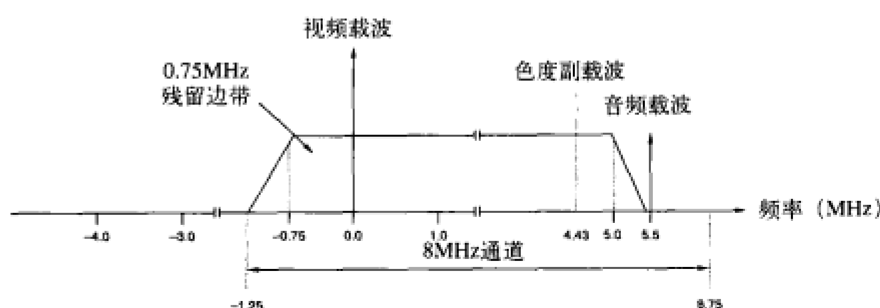
B、G	75%
D、H、M、N	75%
I	76%

白色电平（%调制）

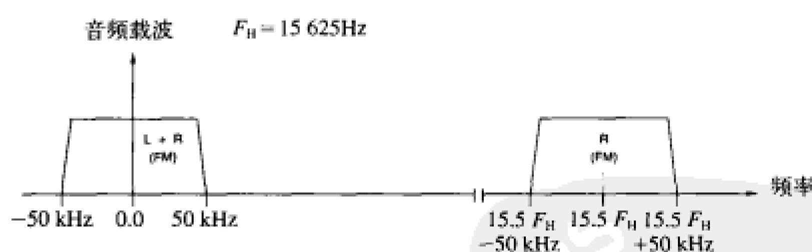
B、G、H、M、N	10%
D	10%
I	20%



（a）基带复合视频的频谱



（b）包含音频信息的典型通道的频谱



（c）Zweiton模拟立体声音频信息的详细频谱

图8-20 （G）PAL制式传输通道

282 ~ 286

注意PAL制式使用一组不同的视频和音频IF频率（单位为MHz）：

	视频	音频	
B、G	38.900	33.400	澳大利亚
B	36.875	31.375	中国
D	37.000	30.500	OIRT
D	38.900	32.400	
I	38.900	32.900	英国
I	39.500	33.500	
M、N	45.750	41.250	

在这点上，音频信息叠加到音频副载波上。如图8-21所示，单声道L＋R音频信号被处理后去驱动FM（Frequency Modulation，频率调制）调制器。然后将FM调制器的输出叠加到IF视频信号中。

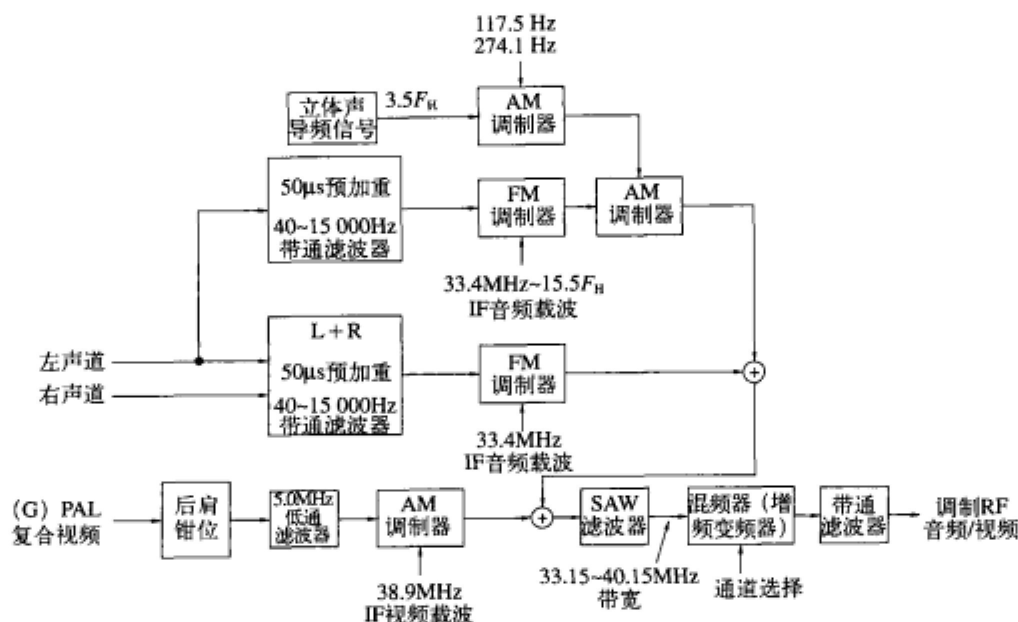


图8-21 (G) PAL的典型RF调制实现：Zweiton立体声音频

作为残留边带滤波器，SAW滤波器滤波IF信号。混频器（或者增频变频器），将IF信号和期望的广播频率混合起来。混频过程中将产生和频与差频，因此使用一个带通滤波器将差频信号抽取出来。

1. 立体声音频（模拟）

图8-21示范了模拟立体声音频的标准（ITU-R BS.707），该标准也称为Zweiton或者A2。L+R信息通过FM副载波传送。R信息，或者辅助L+R音频信号，也是通过FM传送，但它的副载波为$+15.5F_H$。

若立体声或双路单声道信号存在，那么频率为$+15.5F_H$的FM副载波要经过54.6875kHz（$3.5F_H$）的副载波的幅度调制。这一54.6875kHz的副载波或者在117.5Hz（$F_H/133$）处使用50%幅度调制表示立体声音频，或者在274.1Hz（$F_H/57$）处使用50%幅度调制表示双路单声道音频。

使用这种制式的国家包括澳大利亚、奥地利、中国、德国、意大利、马来西亚、荷兰、斯洛文尼亚和瑞士。

287 ↕ 289

2. 立体声音频（数字）

数字立体声音频使用NICAM 728（Near Instantaneous Companded Audio Multiplex，准瞬时压扩音频复用）标准，在BS.707和ETSI EN 300 163中讨论。该标准由BBC和IBA共同开发，以提高伴音质量，提供多个数字伴音或者数据通道，降低传输串挠。

音频副载波或者位于视频载波上面5.85MHz处[对于（B、D、G、H）PAL和（L）SECAM制式]，或者位于视频载波上面6.552MHz处[对于（I）PAL制式]。

使用NICAM 728的国家包括比利时、中国、丹麦、芬兰、法国、匈牙利、新西兰、挪威、新加坡、韩国、西班牙、瑞典和英国。

NICAM 728是一种使用32kHz采样频率和14bit解析度的数字制式。它使用的是728 kbit/s的速率，所以命名为NICAM 728。NICAM 728数据以帧传输，每帧包含1ms音频。如图8-22所示，每个帧包括：

- 8bit帧同步字（01001110）
- 5bit控制信息码（C0~C4）
- 11bit未定义数据（AD0~AD10）
- 704bit音频数据（A000~A703）

帧同步字	控制信息位	附加数据位	704 bit音频数据
0, 1, 0, 0, 1, 1, 1, 0,	C0, C1, C2, C3, C4,	AD0, AD1, AD2, AD3, AD4, AD5, AD6, AD7, AD8, AD9, AD10,	A000, A044, A088, …A660, A001, A045, A089, …A661, A002, A046, A090, …A662, A003, A047, A091, …A663, ⋮ A043, A087, A131, …A703

图8-22 NICAM 728每帧中的比特流

对于前8个连续帧，C0是1，对于后8个连续帧，C0是0，定义了一个有16帧的帧序列。C1~C3指定所传送音频的格式：000表示一路单立体声信号，左通道样本编号为奇数，右通道样本编号为偶数；010表示以交变帧传送两路独立的单声道音频；100表示以交变帧传送一路单声道通道和一路352kbit/s的数据通道；110表示一路704kbit/s的数据通道。如果模拟伴音和数字伴音相同，C4是1。

290 ~ 292

- 立体声音频编码

每个通道上的32个14bit样本（1ms音频、以2的补码进行编码）根据ITU-T J.17曲线进行预加重处理。

32个样本中的最大正或者负样本用来确定要传送全部32个样本所使用的10位。每个通道上的3个幅度位（range bit）（$R0_L$、$R1_L$、$R2_L$和$R0_R$、$R1_R$、$R2_R$）用来指定压缩系数。D13是符号位（0表示正）。

D13~D0	R2~R0	使用位
01xxxxxxxxxxxx	111	D13, D12~D4
001xxxxxxxxxxx	110	D13, D11~D3
0001xxxxxxxxxx	101	D13, D10~D2
00001xxxxxxxxx	011	D13, D9~D1
000001xxxxxxxx	101	D13, D8~D0
0000001xxxxxxx	010	D13, D8~D0
0000000xxxxxxx	00x	D13, D8~D0
1111111xxxxxxx	00x	D13, D8~D0
1111110xxxxxxx	010	D13, D8~D0
111110xxxxxxxx	100	D13, D8~D0
11110xxxxxxxxx	011	D13, D9~D1
1110xxxxxxxxxx	101	D13, D10~D2
110xxxxxxxxxxx	110	D13, D11~D3
10xxxxxxxxxxxx	111	D13, D12~D4

每个样本的6个MSB添加一个奇偶校验位，从而每个样本有11 bit。64个样本相互交织，生成的结果就是L0、R0、L1、R1、L2、R2、…、L31、R31，编号为0~63。

奇偶校验位用来向解码器传递每个通道使用的压缩系数［“奇偶校验信令（signalling-in-parity）”］。

如果$R2_L$＝0，使用样本0、6、12、18、…、48的偶校验。如果$R2_L$＝1，使用奇校验。

如果$R2_R$＝0，使用样本1、7、13、19、…、49的偶校验。如果$R2_R$＝1，使用奇校验。

如果$R1_L=0$，使用样本2、8、14、20、…、50的偶校验。如果$R1_L=1$，使用奇校验。

如果$R1_R=0$，使用样本3、9、15、21、…、51的偶校验。如果$R1_R=1$，使用奇校验。

如果$R0_L=0$，使用样本4、10、16、22、…、52的偶校验。如果$R0_L=1$，使用奇校验。

如果$R0_R=0$，使用样本5、11、17、23、…、53的偶校验。如果$R0_R=1$，使用奇校验。

样本54~63的奇偶校验通常采用偶校验。但是，可以修改它们，从而可以传输两个额外的信息位：

293

如果CIB0＝0，使用样本54、55、56、57和58的偶校验。如果CIB0＝1，使用奇校验。

如果CIB1＝0，使用样本59、60、61、62和63的偶校验。如果CIB1＝1，使用奇校验。

音频数据进行了位交织处理，如图8-22所示，以减少传输过程中突发误码的影响。如果各个位按0~703编号，那么它们按照0、44、88、…、660、1、45、89、…、661、2、46、90、…、703顺序传送。

除了帧同步字外，整个帧和一个1bit的伪随机二进制序列（Pseudo-Random Binary Sequence，PRBS）进行异或运算。PRBS发生器在每个帧的帧同步字之后重新被初始化，从而伪随机二进制序列的第一个位处理的是C0位。PRBS的生成多项式为x^9+x^4+1，初始值为111111111。

实际传输过程包括将728kbit/s比特流形成比特对，然后使用差分正交相移键控（Differential Quadrature Phase-Shift Keying，DQPSK）每秒生成356k符号。如果符号是00，那么副载波相位保持不变。如果符号是01，那么副载波相位延迟90°。如果符号是11，那么副载波相位倒相。如果符号是10，那么副载波相位超前90°。

最后，符号进行频谱成形，形成一个－30dB的约700kHz[（I）PAL制式]或者约500kHz［（B、G）PAL制式］带宽。

❑ 立体声音频解码

PLL加锁NICAM副载波频率，恢复表示编码符号的相位变化。解码的符号生成728kbit/s的比特流。

接着查找帧同步字，对剩余的位和一个本地生成的PRBS进行异或，恢复包。然后为前8帧和后8帧测试C0位，验证是不是NICAM 728比特流。

接下来对位交织处理的音频数据进行去位交织，解出"奇偶校验信令"：

大多数表决发生在样本0、6、12、…、48的奇偶校验上。如果是偶校验，$R2_L=0$；如果是奇校验，$R2_L=1$。

大多数表决发生在样本1、7、13、…、49的奇偶校验上。如果是偶校验，$R2_R=0$；如果是奇校验，$R2_R=1$。

大多数表决发生在样本2、8、14、…、50的奇偶校验上。如果是偶校验，$R1_L=0$；如果是奇校验，$R1_L=1$。

大多数表决发生在样本3、9、15、…、51的奇偶校验上。如果是偶校验，$R1_R=0$；如果是奇校验，$R1_R=1$。

大多数表决发生在样本4、10、16、…、52的奇偶校验上。如果是偶校验，$R0_L=0$；如果是奇校验，$R0_L=1$。

294

大多数表决发生在样本5、11、17、…、53的奇偶校验上。如果是偶校验，$R0_R=0$；如果是奇校验，$R0_R=1$。

大多数表决发生在样本54、55、56、57和58的奇偶校验上。如果是偶校验，CIB0＝0；如果是奇校验，CIB0＝1。

大多数表决发生在样本59、60、61、62和63的奇偶校验上。如果是偶校验，CIB1＝0；如果

是奇校验，CIB1＝1。

任何奇偶校验与表决不符的样本都将忽略，替换为一个内插值。

左通道使用幅度位$R2_L$、$R1_L$和$R0_L$确定在编码过程中丢弃哪些位于符号位之下的位。将符号位复制到这些位置上生成一个14 bit的样本。

右通道处理过程相似，使用幅度位$R2_R$、$R1_R$和$R0_R$。然后使用J.17曲线对两个通道进行去加重（de-emphasize）处理。

❑ 双路单声道音频编码

两个32个14bit样本（2ms音频，以2的补码进行编码）的块根据ITU-T J.17规范进行预加重处理。和立体声音频一样，每个块有3个幅度位（$R0_A$、$R1_A$、$R2_A$和$R0_B$、$R1_B$、$R2_B$），用来表示压缩系数。与立体声音频不同，样本没有进行交织处理。

如果$R2_A$＝0，使用样本0、3、6、9、…、24的偶校验。如果$R2_A$＝1，使用奇校验。

如果$R2_B$＝0，使用样本27、30、33、…、51的偶校验。如果$R2_B$＝1，使用奇校验。

如果$R1_A$＝0，使用样本1、4、7、10、…、25的偶校验。如果$R1_A$＝1，使用奇校验。

如果$R1_B$＝0，使用样本28、31、34、…、52的偶校验。如果$R1_B$＝1，使用奇校验。

如果$R0_A$＝0，使用样本2、5、8、11、…、26的偶校验。如果$R0_A$＝1，使用奇校验。

如果$R0_B$＝0，使用样本29、32、35、…、53的偶校验。如果$R0_B$＝1，使用奇校验。

音频数据进行了位交织处理；但是，奇数包包含音频通道1的64个样本，而偶数包包含音频通道2的64个样本。其他处理过程与立体声音频的处理过程相同。

8.2.7 模拟频道分配

表8-5至表8-7列举了各种PAL制式的VHF、UHF及有线的频道分配情况。

表8-5 各个国家（B、I）PAL制的模拟广播和有线电视标称频率

频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
（B）PAL，澳大利亚，7MHz频道				（B）PAL，意大利，7MHz频道			
0	46.25	51.75	45~52	A	53.75	59.25	52.5~59.5
1	57.25	62.75	56~63	B	62.25	67.75	61~68
2	64.25	69.75	63~70	C	82.25	87.75	81~88
3	86.25	91.75	85~92	D	175.25	180.75	174~181
4	95.25	100.75	94~101	E	183.75	189.25	182.5~189.5
5	102.25	107.75	101~108	F	192.25	197.75	191~198
5A	138.25	143.75	137~144	G	201.25	206.75	200~207
6	175.25	180.75	174~181	H	210.25	215.75	209~216
7	182.25	187.75	181~188	H-1	217.25	222.75	216~223
8	189.25	194.75	188~195	H-2	224.25	229.75	223~230
9	196.25	201.75	195~202	–	–	–	–
10	209.25	214.75	208~215	–	–	–	–
11	216.25	221.75	215~222	–	–	–	–
12	223.25			–	–	–	–
（I）PAL，爱尔兰，8MHz频道				（B）PAL，新西兰，7MHz频道			
1	45.75	51.75	44.5~52.5	1	45.25	50.75	44~51
2	53.75	59.75	52.5~60.5	2	55.25	60.75	54~61

（续）

频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
（I）PAL，爱尔兰，8MHz频道				（B）PAL，新西兰，7MHz频道			
3	61.75	67.75	60.5~68.5	3	62.25	67.75	61~68
4	175.25	181.25	174~182	4	175.25	180.75	174~181
5	183.25	189.25	182~190	5	182.25	187.75	181~188
6	191.25	197.25	190~198	6	189.25	194.75	188~195
7	199.25	205.25	198~206	7	196.25	201.75	195~202
8	207.25	213.25	206~214	8	203.25	208.75	202~209
9	215.25	221.25	214~222	9	210.25	215.75	209~216

表8-6 [1]英国、[1]爱尔兰、[1]韩国、[1]香港和[2]西欧的模拟广播标称频率

广播频道	视频载波（MHz）	音频载波（MHz）		频道范围（MHz）
		（G、H）PAL	（I）PAL	
2[1]	45.75	51.25	51.75	44.5~52.5
3[1]	53.75	59.25	59.75	52.5~60.5
4[1]	61.75	67.25	67.75	60.5~68.5
5[1]	175.25	180.75	181.25	174~182
6[1]	183.25	188.75	189.25	182~190
7[1]	191.25	196.75	197.25	190~198
8[1]	199.25	204.75	205.25	198~206
9[1]	207.25	212.75	213.25	206~214
10[1]	215.25	220.75	221.25	214~222
2[2]	48.25	53.75	–	47~54
3[2]	55.25	60.75	–	54~61
4[2]	62.25	67.75	–	61~68
5[2]	175.25	180.75	–	174~181
6[2]	182.25	187.75	–	181~188
7[2]	189.25	194.75	–	188~195
8[2]	196.25	201.75	–	195~202
9[2]	203.25	208.75	–	202~209
10[2]	210.25	215.75	–	209~216
11[2]	217.25	222.75	–	216~223
12[2]	224.25	229.75	–	223~230
21	471.25	476.75	477.25	470~478
22	479.25	484.75	485.25	478~486
23	487.25	492.75	493.25	486~494
24	495.25	500.75	501.25	494~502
25	503.25	508.75	509.25	502~510
26	511.25	516.75	517.25	510~518
27	519.25	524.75	525.25	518~526
28	527.25	532.75	533.25	526~534
29	535.25	540.75	541.25	534~542
30	543.25	548.75	549.25	542~550
31	551.25	556.75	557.25	550~558

（续）

广播频道	视频载波（MHz）	音频载波（MHz）		频道范围（MHz）
		（G、H）PAL	（I）PAL	
32	559.25	564.75	565.25	558~566
33	567.25	572.75	573.25	566~574
34	575.25	580.75	581.25	574~582
35	583.25	588.75	589.25	582~590
36	591.25	596.75	597.25	590~598
37	599.25	604.75	605.25	598~606
38	607.25	612.75	613.25	606~614
39	615.25	620.75	621.25	614~622
40	623.25	628.75	629.25	622~630
41	631.25	636.75	637.25	630~638
42	639.25	644.75	645.25	638~646
43	647.25	652.75	653.25	646~654
44	655.25	660.75	661.25	654~662
45	663.25	668.75	669.25	662~670
46	671.25	676.75	677.25	670~678
47	679.25	684.75	685.25	678~686
48	687.25	692.75	693.25	686~694
49	695.25	700.75	701.25	694~702
50	703.25	708.75	709.25	702~710
51	711.25	716.75	717.25	710~718
52	719.25	724.75	725.25	718~726
53	727.25	732.75	733.25	726~734
54	735.25	740.75	741.25	734~742
55	743.25	748.75	749.25	742~750
56	751.25	756.75	757.25	750~758
57	759.25	764.75	765.25	758~766
58	767.25	772.75	773.25	766~774
59	775.25	780.75	781.25	774~782
60	783.25	788.75	789.25	782~790
61	791.25	796.75	797.25	790~798
62	799.25	804.75	805.25	798~806
63	807.25	812.75	813.25	806~814
64	815.25	820.75	821.25	814~822
65	823.25	828.75	829.25	822~830
66	831.25	836.75	837.25	830~838
67	839.25	844.75	845.25	838~846
68	847.25	852.75	853.25	846~854
69	855.25	860.75	861.25	854~862

表8-7 英国、爱尔兰、韩国、香港和西欧的模拟有线电视标称频率

有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）	有线频道	视频载波（MHz）	音频载波（MHz）	频道范围（MHz）
E2	48.25	53.75	47~54	S01	69.25	74.75	68~75
E3	55.25	60.75	54~61	S02	76.25	81.75	75~82
E4	62.25	67.75	61~68	S03	83.25	88.75	82~89

（续）

有线频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)	有线频道	视频载波 (MHz)	音频载波 (MHz)	频道范围 (MHz)
S11	231.25	236.75	230~237	S14	252.25	257.75	251~258
S12	238.25	243.75	237~244	S15	259.25	264.75	258~265
S13	245.25	250.75	244~251	S16	266.25	271.75	265~272
S1	105.25	110.75	104~111	S17	273.25	278.75	272~279
S2	112.25	117.75	111~118	S18	280.25	285.75	279~286
S3	119.25	124.75	118~125	S19	287.25	292.75	286~293
S4	126.25	131.75	125~132	S20	294.25	299.75	293~300
S5	133.25	138.75	132~139	S21	303.25	308.75	302~310
S6	140.75	145.75	139~146	S22	311.25	316.75	310~318
S7	147.75	152.75	146~153	S23	319.25	324.75	318~326
S8	154.75	159.75	153~160	S24	327.25	332.75	326~334
S9	161.25	166.75	160~167	S25	335.25	340.75	334~342
S10	168.25	173.75	167~174	S26	343.25	348.75	342~350
–	–	–	–	S27	351.25	356.75	350~358
–	–	–	–	S28	359.25	364.75	358~366
–	–	–	–	S29	367.25	372.75	366~374
E5	175.25	180.75	174~181	S30	375.25	380.75	374~382
E6	182.25	187.75	181~188	S31	383.25	388.75	382~390
E7	189.25	194.75	188~195	S32	391.25	396.75	390~398
E8	196.25	201.75	195~202	S33	399.25	404.75	398~406
E9	203.25	208.75	202~209	S34	407.25	412.75	406~414
E10	210.25	215.75	209~216	S35	415.25	420.75	414~422
E11	217.25	222.75	216~223	S36	423.25	428.75	422~430
E12	224.25	229.75	223~230	S37	431.25	436.75	430~438
–	–	–	–	S38	439.25	444.75	438~446
–	–	–	–	S39	447.25	452.75	446~454
–	–	–	–	S40	455.25	460.75	454~462
–	–	–	–	S41	463.25	468.75	462~470

注意有线电视系统按惯例重新分配了频道号，以交变频率，从而将干扰最小化并提供多种级别的节目（例如两种版本的收费电影频道：其中一个针对订户，另一个针对预演阶段的非订户）。

8.2.8　亮度方程推导

由RGB信息生成亮度的方程由接收机所使用的三基色色度和真正的白光源决定。

RGB三基色色度和基准白光（CIE照明D_{65}）为：

R：$x_r=0.64$　　$y_r=0.33$　　$z_r=0.03$

G：$x_g=0.29$　　$y_g=0.60$　　$z_g=0.11$

B：$x_b=0.15$　　$y_b=0.06$　　$z_b=0.79$

295　白色：$x_w=0.3127$　　$y_w=0.3290$　　$z_w=0.3583$

其中x和y由CIE 1931色度坐标指定，而z由$x+y+z=1$计算。

和NTSC一样，将已知数据代入上式，求解出K_r、K_g和K_b：

$$\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix} = \begin{bmatrix} 0.3127/0.3290 \\ 1 \\ 0.3583/0.3290 \end{bmatrix} \begin{bmatrix} 0.64 & 0.29 & 0.15 \\ 0.33 & 0.60 & 0.06 \\ 0.03 & 0.11 & 0.79 \end{bmatrix}^{-1} = \begin{bmatrix} 0.674 \\ 1.177 \\ 1.190 \end{bmatrix}$$

Y定义为：

$$\begin{aligned} Y &= (K_r y_r)R' + (K_g y_g)G' + (K_b y_b)B' \\ &= (0.674)(0.33)R' + (1.177)(0.60)G' \\ &\quad + (1.190)(0.06)B' \end{aligned}$$

或者：

$$Y = 0.222R' + 0.706G' + 0.071B'$$

但是，仍然可以使用标准的$Y = 0.299R' + 0.587G' + 0.114B'$方程式。为了减少色彩误差，需要对接收机进行调整。

8.2.9 PALplus

PALplus（ITU-R BT.1197和ETSI ETS 300 731）是一个从1990年开始的协作项目的成果，该项目由几个欧洲广播商承担。在1995年时，他们希望提供一种与现有接收机兼容的增强清晰度电视（Enhanced Definition Television，EDTV）。从1994年开始，就有少数几个广播商开始传送PALplus。

PALplus图像是16：9宽高比的画面。在传统电视机上，它显示为一个拥有430有效行的16：9的Letterbox图像。在PALplus电视机上，它显示为一个拥有574有效行的16：9的图像，而且垂直解析度还得到增强。亮度信息可以使用完整的视频带宽。串色干扰可以通过干净编码（clean encoding）得以减少。

1. 宽屏信令

第23行包含一个宽屏信令（Wide Screen Signaling，WSS）控制信号，该信令由ITU-R BT.1119和ETSI EN 300 294定义，由PALplus电视机使用。该信令含义如下。

节目宽高比：

- 完全模式4：3
- Letterbox模式14：9居中
- Letterbox模式14：9置顶
- 完全模式14：9居中
- Letterbox模式16：9居中
- Letterbox模式16：9置顶
- 完全模式16：9变形
- Letterbox模式>16：9居中

增强服务：

- 摄像机模式
- 影片模式

字幕：

- 图文电视字幕（Teletext Subtitle）呈现
- 开放字幕（Open Subtitle）呈现

PALplus方式定义为16：9 Letterbox居中模式、摄像机模式或者影片模式，通过调制提供辅助信号并使用干净编码技术。图文电视字幕可以呈现，也可以不呈现，而开放字幕只能在有效图像区域内呈现。

296~300

在PALplus信号传输过程中，在编码之前，任何位于第23行和第623行上的有效视频都将被清除。除WSS数据之外，第23行还包含48±1周期的300±9mV副载波和−U相位，从0_H之后的51μs±250ns开始。第623行包含一个10μs±250ns的白色脉冲，从0_H之后的20μs±250ns开始。

PALplus电视机可以对影片模式信号去进行隔行处理并在一个50Hz逐行扫描显示器上显示，也可以使用场重复方法在一个100Hz的隔行显示器上显示。

2. 重影消除

第318行上还有一个可选的重影消除（ghost cancellation）信号，该信号由ITU-R BT.1124和ETSI ETS 300 732定义，允许支持这一功能的电视机测量重影信号，消除有效视频中的重影。PALplus电视机可以支持也可以不支持这一功能。

3. 垂直滤波

所有的PALplus源最初都是一个16：9的YCbCr变形图像（anamorphic image），该图像占用全部576个有效扫描行。第23行和第623行上的所有有效视频在编码之前都将被清除（因为这些行用于WSS和参考信息），从而使每帧的有效行只有574行。第24~310行和第336~622行用于有效视频。

在传输之前，16：9图像的574行有效扫描行压缩到430行扫描行中。为了避免图像的混叠问题，垂直解析度通过低通滤波降低。

对于Y信号而言，通过使用正交镜像滤波器（Quadrature Mirror Filter，QMF）高通和低通滤波对实现垂直滤波。通过应用QMF处理过程，可以对高通和低通滤波信息重新采样、传送，然后以最小的损耗重组。

Y信号的QMF低通滤波输出重新采样为原始高度的3/4；混叠问题对信息所造成的损失极少。在干净编码之后，它就成为传统4：3电视机显示的Letterbox模式信号。

Y信号的QMF高通滤波输出包含原始场频的其余部分。它用来生成没有被Letterbox模式图像使用的黑场扫描行传送的辅助信号。

❑ 影片模式

影片模式广播拥有来源于同一个图像的帧的两个场，就像在电视电影上显示电影一样。

在影片模式中，每帧的最大垂直解析度为每个有效图像高度287周期（cycle per active picture height，cph），受每帧有效扫描行574行限制。

Y信号的垂直解析度降低为215cph，从而只需使用430行有效扫描行就可以进行传送。QMF低通和高通滤波器将Y信号垂直信息分隔成DC-215cph和216-287cph分量。

Y信号低通信息被重新扫描成430行，变成Letterbox模式图像。因为场频受最大215cph限制，所以没有信息丢失。

Y信号高通输出的行数骤减，因此4行中只用传送1行。这144行用来传送辅助信号。因为QMF处理，所以在行数骤减情况下没有信息丢失。

Letterbox模式图像中间有430行，上面和下面各有72行用来传送144行辅助信号。这样就生成一个标准的574行有效扫描行图像，原始图像以正确的宽高比显示，并且位于辅助信号的中间。

301

包含300mV辅助信号的扫描行使用U副载波调制，因此它们看上去是黑色的，对观众不可见。

在固定ColorPlus处理过程之后，574行扫描行经过PAL编码，作为标准的隔行PAL帧传送。

❑ 摄像机模式

摄像机（或者视频）模式假定帧的各个场之间相互独立，就像摄像机拍摄一个运动中的场景

一样。因此，图像在场之间可能有变化。摄像机模式只进行场内的处理。

在摄像机模式中，每场的最大垂直解析度大约是每个有效图像高度143周期（cph），它受每场有效扫描行287行的限制。

Y信号的垂直解析度降低为107cph，从而只需使用215行有效行就可以进行传送。QMF低通和高通滤波器对将Y信号垂直信息分隔成DC-107cph和108-143cph分量。

Y信号低通信息被重新扫描成215行，变成Letterbox模式图像。因为场频受最大107cph限制，所以没有信息丢失。

Y信号高通输出的行数骤减，因此4行中只用传送1行。这72行用来传输辅助信号。因为QMF处理，所以在行数骤减情况下没有信息丢失。

Letterbox模式图像中间有215行，上面和下面各有36行用来传输72行辅助信号。这样就生成一个287行有效行图像，原始图像以正确的宽高比显示，并且位于辅助信号的中间。包含300mV辅助信号的扫描行使用U副载波调制，因此它们看上去是黑色的，对观众不可见。

在固定或者运动自适应ColorPlus处理之后，287行扫描行经过PAL编码，作为一个PAL场传送。

4. 干净编码

PALplus信号中只有Letterbox模式这一部分信号经过干净编码处理。辅助信号并不是实际的PAL视频。但是，它们太接近于视频信号，从而能够穿过传输路径，只是在标准电视机上仍然保持相对不可见。

5. ColorPlus处理

❑ 固定ColorPlus

影片模式使用固定ColorPlus技术，从而利用帧的两个场之间没有运动这种情况。

固定ColorPlus处理依赖于处于相反相位的复合PAL信号在分离312行时的副载波相位。如果这两行拥有相同的亮度和色度信息，就可以通过相互之间添加和抽取复合信号来分离。添加复合信号意味着取消色度信号，剩下亮度信号。抽取复合信号意味着取消亮度信号，剩下色度信号。

在实际中，位于3MHz之上的Y信息（Y_{HF}）是帧内平均的，因为它与调制色度共享频谱。对于行[*n*]而言，Y_{HF}的计算如下所示：

对于430行Letterbox模式图像：$0 \leqslant n \leqslant 214$

$$Y_{HF(60+n)} = 0.5[Y_{HF(372+n)} + Y_{HF(60+n)}]$$

$$Y_{HF(372+n)} = Y_{HF(60+n)}$$

接下来Y_{HF}添加到低频Y（Y_{LF}）信息上。Cb和Cr也采用相同的帧内平均过程。然后430行Letterbox模式图像进行PAL编码。

302

因此，3MHz以上的Y信息和CbCr信息，在行[*n*]和行[*n*+312]上是相同的。低于3MHz的Y信息在行[*n*]和行[*n*+312]上可能不同。为了辅助信号，解码器需要重建287cph的完整垂直解析度。

❑ 运动自适应ColorPlus（MACP）

摄像机模式根据场间的运动量，或者使用运动自适应ColorPlus，或者使用固定ColorPlus。这就要求编码器和解码器都提供运动检测器（motion detector）。

为了检测移动，需要比较行[*n*]和行[*n*+312]上的CbCr数据。如果它们匹配，那么假定没有运动，使用固定ColorPlus运算。如果CbCr数据不匹配，那么假定存在动动，使用运动自适应ColorPlus运算。

在运动自适应ColorPlus运算过程中，添加到Y_{LF}上的Y_{HF}量取决于$CbCr_{(n)}$和$CbCr_{(n+312)}$之间的差值。对于最大的CbCr差值，行[*n*]和行[*n*+312]上不传输Y_{HF}数据。

另外，帧内平均的CbCr数据量与直接CbCr数据的混合取决于CbCr(*n*)和$CbCr_{(n+312)}$之间的差值。

对于最大的CbCr差值，行[*n*]和行[*n*+312]上只传送直接的CbCr数据。

8.3 SECAM概述

SECAM（Sequentiel Couleur Avec Mémoire或者Sequential Color with Memory，按序传送彩色与存储）制式是法国开发的，于1967年开始广播。该制式采用这样的思想，即既然可以在水平方向上限制彩色的带宽，为什么不能在垂直方向上限制彩色的带宽？添加到黑白信号上的两部分彩色信息（Db和Dr）能够在正交的行上传送，从而避免可能的互串。

SECAM制式的接收机要求提供存储器，存储一行扫描行，和下一行扫描行一起使用，另外SECAM制式接收机还要求使用行转换识别技术。

和PAL制式一样，SECAM制式也是625行、50场/秒、2：1隔行制式。SECAM制式曾被法国以外的其他国家采用，但是，由于PAL制式专业和家用设备的丰富多样，许多国家开始转向了PAL制式。

8.3.1 亮度信息

由R′G′B′推导的黑白亮度（Y）信号为：

$Y = 0.299R' + 0.587G' + 0.114B'$

和NTSC及PAL制式一样，SECAM制式的亮度信号也占用整个视频带宽。SECAM制式有几种变体，每种变体的视频带宽和音频副载波的位置有所不同。根据不同的SECAM标准，视频信号的带宽有5.0MHz和6.0MHz两种。

8.3.2 彩色信息

SECAM制式在一行上传送Db信息，在下一行上传送Dr信息，亮度信息在每一行上都传送。Db和Dr是加权后的B′−Y和R′−Y：

$Dr = -1.902(R' - Y)$

$Db = 1.505(B' - Y)$

303 因为行的数量是奇数，所以某场上包含Db信息的行，在下一场上包含Dr信息。解码器要求一行的延迟，由Db和Dr开关同步切换，从而为了转换为YCbCr或者RGB，Db和Dr信息能够同时存在。

8.3.3 彩色调制

SECAM使用FM调制来传送Db和Dr色差信息，每个分量都有自己的副载波。

Db和Dr信息使用1.3MHz的低通滤波器滤波，并进行预加重处理。预加重曲线如下所示：

$$A = \frac{1 + j\left(\frac{f}{85}\right)}{1 + j\left(\frac{f}{255}\right)}$$

其中f为信号频率，单位为kHz。

在预加重处理之后，Db和Dr频率分别调制它们相应的副载波。每个副载波的频率定义如下：

$F_{OB} = 272F_H = 4.250000$ MHz (±2kHz)

$F_{OR} = 282F_H = 4.406250$ MHz (±2kHz)

这些频率不表现彩色信息。标称Dr偏差为±280kHz，标称Db偏差为±230kHz。图8-23描述了色差信号的频率调制过程。频偏偏移的选择反映出这样一种思想：使表示关键色彩的频率避开

频谱上限，以最小化失真。

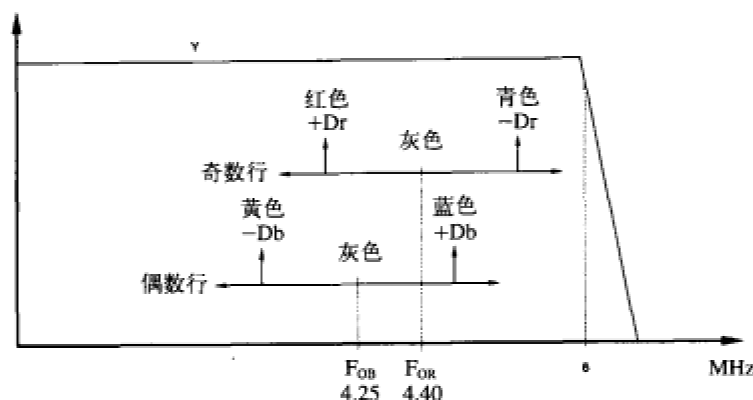


图8-23 SECAM制式FM彩色调制

在Db和Dr信息调制、应用了副载波预加重之后，副载波幅度的改变就成为频率变化的一个函数。这样做的目的是减小低亮度区域内副载波的可视性，提高高饱和彩色的信噪比。该预加重曲线如下所示：

$$G = M\frac{1+\text{j}16F}{1+\text{j}1.26F}$$

其中$F=(f/4286)-(4286/f)$，f是瞬时副载波频率，单位为kHz，$2M$＝亮度幅度的（23±2.5%）。

如图8-24所示，Db和Dr信息分别在正交扫描行上传送。彩色副载波的初始相位也进行了修改，如表8-8所示，以进一步减小副载波可视性。注意SECAM制式中的副载波相位信息没有携带图像信息。

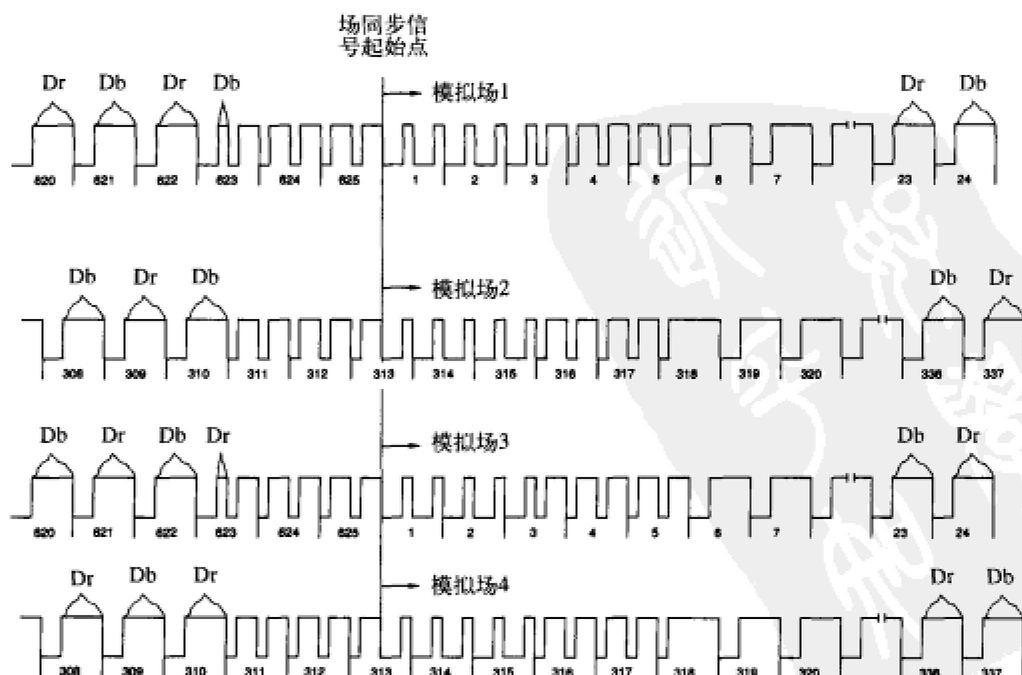


图8-24 4场SECAM序列。均衡和齿脉冲的详细信息请参见图8-5

表8-8　SECAM制式副载波时序

每行中初始相位副载波的变化定义
帧到帧：0°、180°、0°、180°…… 行到行：0°、0°、180°、0°、0°、180°……或者0°、0°、0°、180°、180°、180°……

8.3.4　复合视频的生成

副载波数据加上亮度信息以及合适的水平和场同步信号、消隐信息及色同步信号就生成复合视频。

和PAL制式一样，SECAM制式要求一些标识行切换序列的方法。现代实践通常在大多数水平同步信号之后使用一个F_{OR}/F_{OB}色同步信号，推导出切换同步信息，如图8-25所示。

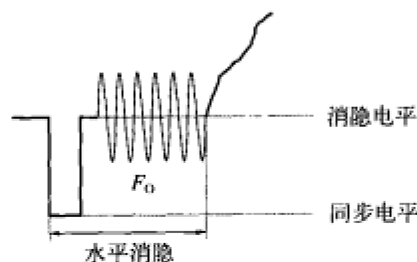


图8-25　SECAM制式色度同步信号

8.3.5　SECAM标准

图8-26示范了几种常用的SECAM制式表示方法。各个字母分别代表各种行频和场频、视频带宽（5.0MHz或6.0MHz）、音频载波相对频率和RF频道带宽的黑白电视标准。SECAM制式指的是向黑白电视信号添加彩色信息的技术。详细的时序参数位于表8-9中。

304

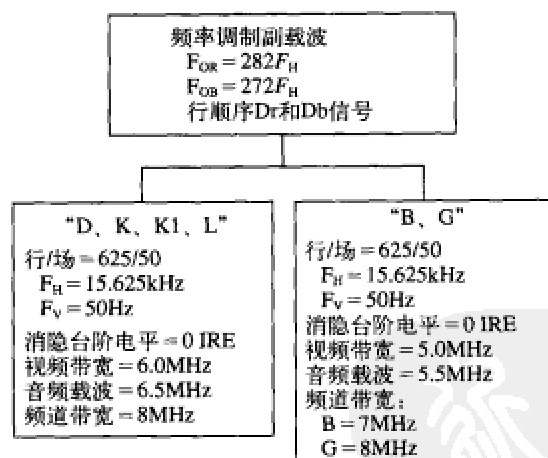


图8-26　常见SECAM制式

表8-9a　彩色视频信号的基本特征

	M	N	B、G	H	I	D、K	K1	L
每帧扫描行数	525	625	625					
场频（场/秒）	59.94	50	50					
行频（Hz）	15 734	15 625	15 625					
峰值白色电平（IRE）	100	100	100					
同步脉冲顶部电平（IRE）	−40	−40（−43）	−43					
消隐台阶电平（IRE）	7.5±2.5	7.5±2.5（0）	0					
峰值视频电平（IRE）	120		133		133	115	115	125
接收机的gamma值	2.2	2.8	2.8	2.8	2.8	2.8	2.8	2.8
视频带宽（MHz）	4.2	5.0（4.2）	5.0	5.0	5.5	6.0	6.0	6.0
亮度信号	Y＝0.299R′＋0.587G′＋0.114B′（RGB经过gamma校正）							

注：括号中的值适用于阿根廷使用的（N_C）PAL制式。

表8-9b 行同步信号的细节

特　征	M	N	B、D、G、H、I K、K1、L、N_C
标称行周期（μs）	63.5555	64	64
行消隐期（μs）	10.7±0.1	10.88±0.64	11.85±0.15
有效视频起始点0_H（μs）	9.2±0.1	9.6±0.64	10.5
前肩（μs）	1.5±0.1	1.92±0.64	1.65±0.15
行同步脉冲（μs）	4.7±0.1	4.99±0.77	4.7±0.2
行消隐的上升和下降时间（10%，90%）（ns）	140±20	300±100	300±100
行同步脉冲的上升和下降时间（10%，90%）（ns）	140±20	≤250	250±50

注：1. 0_H位于场同步信号下降边的50%处。
　　2. 对于拥有不同规范和容限的不同标准，本表列举的是最严格的规范和容限。
　　3. 时序信息是在合适信号边半幅度位置处测量的。

表8-9c 场同步信号的细节

特　征	M	N	B、D、G、H、I K、K1、L、N_C
场周期（ms）	16.6833	20	20
场消隐期	20行	19~25行	25行
场消隐的上升和下降时间（10%，90%）（ns）	140±20	≤250	300±100
均衡和同步序列的持续时间	3行	3行	2.5行
均衡脉冲宽度（μs）	2.3±0.1	2.43±0.13	2.35±0.1
齿脉冲宽度（μs）	4.7±0.1	4.7±0.8	4.7±0.1
同步和均衡脉冲的上升和下降时间（10%，90%）（ns）	140±20	<250	250±50

注：1. 为了应对不同标准拥有不同的规范和容限，本表列举的是最严格的规范和容限。
　　2. 时序信息是在合适信号边半幅度位置处测量的。

表8-9d 彩色视频信号的基本特征

	M/NTSC	M/PAL	B、D、G、H、I、N/PAL	B、D、G、K、K1、K/SECAM
色差信号衰减	U、V、I、Q： 1.3MHz处，<2dB 3.6MHz处，>20dB 或者Q： 0.4MHz处，<2dB 0.5MHz处，<6dB 0.6MHz处，>6dB	1.3MHz处，<2dB 3.6MHz处，>20dB	1.3MHz处，<3dB 4MHz处，>20dB （3.6MHz处，>20dB）	1.3MHz处，<3dB 3.5MHz处，>30dB （在低频预加重处理之前）
0H之后色同步信号起始点（μs）	5.3±0.07	5.8±0.1	5.6±0.1	
色同步信号持续时间（周）	9±1	9±1	10±1（9±1）	
色同步信号峰值幅度	40±1 IRE	42.86±4 IRE	42.86±4 IRE	

注：括号中的值适用于阿根廷使用的（N_C）PAL制式。

8.3.6　亮度方程推导

由RGB信息生成亮度的方程由接收机所使用的三基色色度和真正的白光源决定。

RGB三基色色度和基准白光（CIE照明D_{65}）为：

R：$x_r=0.64$　　$y_r=0.33$　　$z_r=0.03$

G：$x_g=0.29$　　$y_g=0.60$　　$z_g=0.11$

B：$x_b=0.15$　　$y_b=0.06$　　$z_b=0.79$

白色：$x_w=0.3127$　　$y_w=0.3290$　　$z_w=0.3583$

305–307

其中x和y由CIE 1931色度坐标指定；而z由$x+y+z=1$计算。和以往一样，将已知数据代入上式，求解出K_r、K_g和K_b：

$$\begin{bmatrix} K_r \\ K_g \\ K_b \end{bmatrix} = \begin{bmatrix} 0.3127/0.3290 \\ 1 \\ 0.3583/0.3290 \end{bmatrix} \begin{bmatrix} 0.64 & 0.29 & 0.15 \\ 0.33 & 0.60 & 0.06 \\ 0.03 & 0.11 & 0.79 \end{bmatrix}^{-1} = \begin{bmatrix} 0.674 \\ 1.177 \\ 1.190 \end{bmatrix}$$

Y定义为：

$$Y=(K_r y_r)R'+(K_g y_g)G'+(K_b y_b)B'$$
$$=(0.674)(0.33)R'+(1.177)(0.60)G'$$
$$+(1.190)(0.06)B'$$

或者：

$$Y=0.222R'+0.706G'+0.071B'$$

但是，仍然可以使用标准的$Y=0.299R'+0.587G'+0.114B'$方程式。为了减少色彩误差，需要对接收机进行调整。

308

8.4　视频测试信号

目前已经定义了许多视频测试信号的产业标准，以帮助测试编码、解码以及传输路径的相对质量，同时执行校准。注意一些视频测试信号无法通过向编码器提供RGB数据而正确生成，在这种情况下，可以使用YCbCr数据生成。

如果视频标准使用7.5 IRE的台阶电平，那么通常只用于视觉检测的测试信号才使用7.5 IRE的台阶电平。用于测试目的的测试信号通常使用0 IRE的台阶电平，从而可以定义一个已知的消隐电平。

8.4.1　彩条概述

彩条是一种标准的视频测试信号。根据视频标准和应用的不同，彩条有多种变体。由于这一原因，本节回顾一下大部分常见的彩条格式。彩条有两个主要的特征：幅度和饱和度。

彩条信号的幅度可如下计算：

$$\text{幅度（\%）}=\frac{\max(R,G,B)_a}{\max(R,G,B)_b}\times 100$$

其中$\max(R,G,B)_a$是彩条中R'G'B'的最大值，$\max(R,G,B)_b$是基准白色中R'G'B'的最大值。

如果R′G′B′某个分量的最小值不是0，那么彩条信号的饱和度就不是100%。饱和度可如下计算：

$$饱和度（\%）=\left[1-\left(\frac{\min(R,G,B)}{\max(R,G,B)}\right)^{\gamma}\right]\times 100$$

309
↕
312

其中min (*R*, *G*, *B*) 和max (*R*, *G*, *B*) 分别是彩条中R′G′B′的最小值和最大值，γ是gamma分量，典型值为[1/0.45]。

1. NTSC彩条

1953年，模拟R′G′B′信号常见的做法是使用一个7.5 IRE的台阶电平，最初的NTSC方程式假定编码器的输入也是这一形式。现在，数字R′G′B′或YCbCr信号通常不包含7.5 IRE的台阶电平，而在编码器内部添加7.5 IRE的台阶电平。

不同的彩条信号由4个幅度描述，分别以百分比表示，用斜线隔开。在该表示法中，暗含着饱和度为100%，所以不指定饱和度。第一个和第二个数字分别为白色幅度和黑色幅度。第三个和第四个数字是导出彩条的白色幅度和黑色幅度。

例如100/7.5/75/7.5彩条，它是7.5%台阶电平的75%彩条，白条设置为100%，黑条设置为7.5%。因为NTSC制式通常使用7.5%台阶电平，所以两种常见的彩条为75/7.5/75/7.5和100/7.5/100/7.5，通常分别简称为75%彩条和100%彩条。75%彩条是最常使用的彩条。电视发送机不会使用超出120 IRE的幅度传送信息。因此，75%彩条用于传输测试。100%彩条可以用于设备之间可能存在直接连接的情况下的测试。75/7.5/75/7.5彩条属于电子工业协会（Electronic Industries Association，EIA）EIA-189-A编码彩条标准。

图8-27示范了一个满屏75% NTSC彩条的典型矢量示波器显示图。图8-28描述了75%彩条的视频波形。

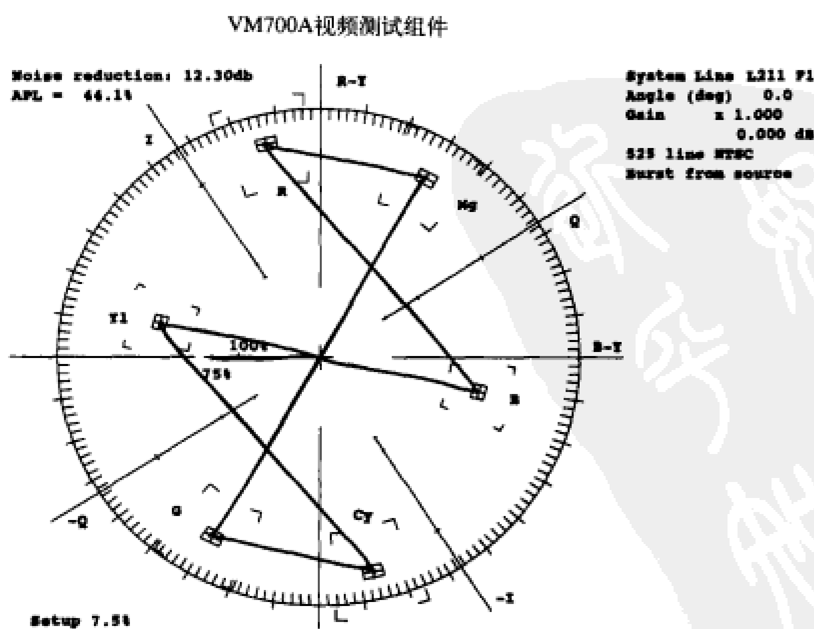


图8-27 75% NTSC彩条的典型矢量示波器显示图

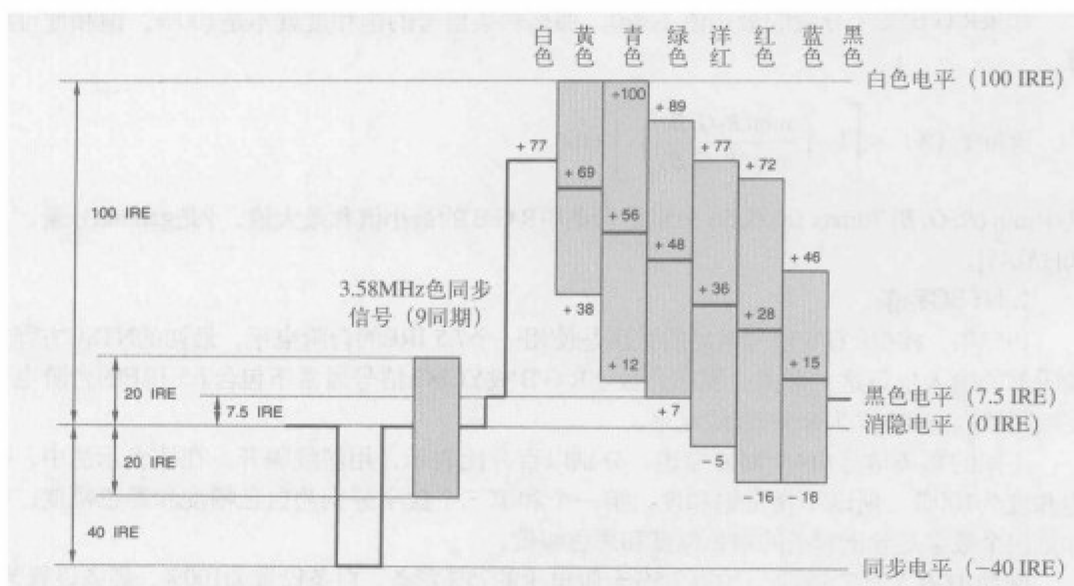


图8-28　75% NTSC彩条的IRE值

表8-10和表8-11分别列举了NTSC两种常见彩条格式的亮度和色度电平。

表8-10　75/7.5/75/7.5（75%）NTSC彩条

	亮度（IRE）	色度电平（IRE）	最小色度偏移（IRE）	最大色度偏移（IRE）	色度相位（度）
白色	76.9	0	–	–	–
黄色	69.0	62.1	37.9	100.0	167.1
青色	56.1	87.7	12.3	100.0	283.5
绿色	48.2	81.9	7.3	89.2	240.7
洋红	36.2	81.9	−4.8	77.1	60.7
红色	28.2	87.7	−15.6	72.1	103.5
蓝色	15.4	62.1	−15.6	46.4	347.1
黑色	7.5	0	–	–	–

表8-11　100/7.5/100/7.5（100%）NTSC彩条

	亮度（IRE）	色度电平（IRE）	最小色度偏移（IRE）	最大色度偏移（IRE）	色度相位（度）
白色	100.0	0	–	–	–
黄色	89.5	82.8	48.1	130.8	167.1
青色	72.3	117.0	13.9	130.8	283.5
绿色	61.8	109.2	7.2	116.4	240.7
洋红	45.7	109.2	−8.9	100.3	60.7
红色	35.2	117.0	−23.3	93.6	103.5
蓝色	18.0	82.8	−23.3	59.4	347.1
黑色	7.5	0	–	–	–

为便于参考，表8-12和表8-13列举了生成标准NTSC彩条的RGB和YCbCr值。假定RGB值的范围为0~255，假定YCbCr值的范围是Y为16~235，Cb和Cr为16~240。另外还假定编码器内提供有

7.5 IRE的台阶电平。

表8-12 75% NTSC彩条的RGB和YCbCr值

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
gamma校正的RGB（gamma＝1/0.45）								
R′	191	191	0	0	191	191	0	0
G′	191	191	191	191	0	0	0	0
B′	191	0	191	0	191	0	191	0
线性RGB								
R	135	135	0	0	135	135	0	0
G	135	135	135	135	0	0	0	0
B	135	0	135	0	135	0	135	0
YCbCr								
Y	180	162	131	112	84	65	35	16
Cb	128	44	156	72	184	100	212	128
Cr	128	142	44	58	198	212	114	128

表8-13 100% NTSC彩条的RGB和YCbCr值

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
gamma校正的RGB（gamma＝1/0.45）								
R′	255	255	0	0	255	255	0	0
G′	255	255	255	255	0	0	0	0
B′	255	0	255	0	255	0	255	0
线性RGB								
R	255	255	0	0	255	255	0	0
G	255	255	255	255	0	0	0	0
B	255	0	255	0	255	0	255	0
YCbCr								
Y	235	210	170	145	106	81	41	16
Cb	128	16	166	54	202	90	240	128
Cr	128	146	16	34	222	240	110	128

2. PAL彩条

与NTSC不同，PAL不支持7.5 IRE的台阶电平，黑色电平和消隐电平相同。不同的彩条信号通常由4个幅度描述，分别以百分比表示，用斜线隔开。第一个和第二个数字分别是非彩条R′G′B′值的最大百分比和最小百分比。第三个和第四个数字分别是彩条R′G′B′值的最大百分比和最小百分比。

因为PAL制式的台阶电平为0%，所以PAL制式两种常见的彩条100/0/75/0和100/0/100/0通常分别简称为75%彩条和100%彩条。75%彩条用于传输测试。100%彩条可以用于设备之间可能存在直接连接的情况下的测试。

100/0/75/0彩条还称为EBU（European Broadcast Union，欧洲广播联盟）彩条。本节中讨论的所有彩条都属于独立电视管理局（Independent Television Authority，ITA）和英国广播公司（British Broadcasting Corporation，BBC）出版的*Specification of Television Standards for 625-line System-I Transmissions*（*1971*）以及*ITU-R BT.471*。 313

图8-29描述了75%彩条的视频波形。图8-30示范了一个满屏75% PAL彩条的典型矢量示波器显

示图。

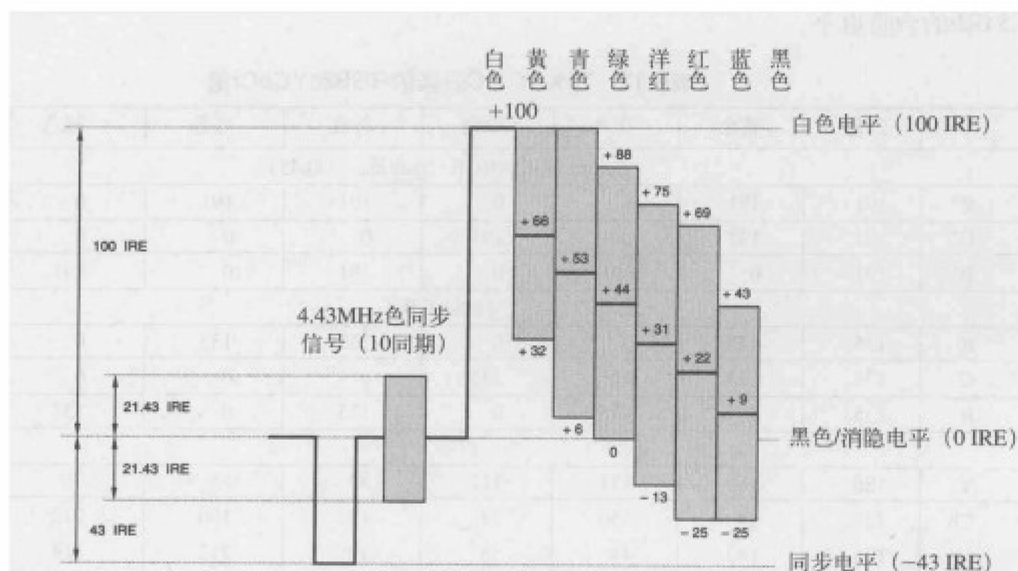


图8-29 75% PAL彩条的IRE值

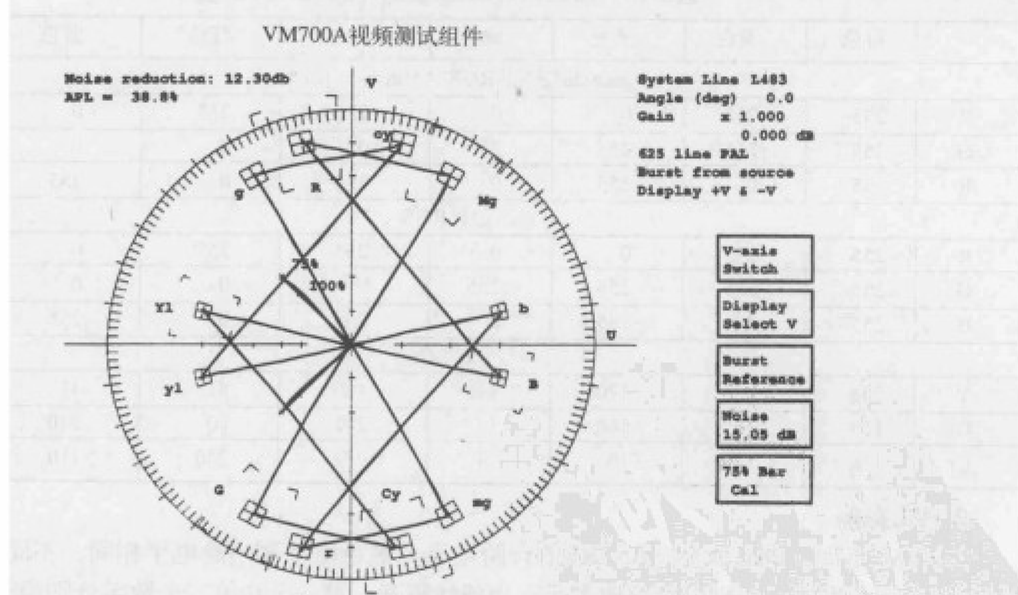


图8-30 75% PAL彩条的典型矢量示波器显示图

表8-14、表8-15和表8-16分别列举了PAL三种常见彩条格式的亮度和色度电平。

表8-14 100/0/75/0（75%）PAL彩条

	亮度（V）	峰－峰色度			色度相位（度）	
		U色轴（V）	V色轴（V）	总计（V）	行n（135°色同步信号）	行（n+1）（225°色同步信号）
白色	0.700	0	–	–	–	–
黄色	0.465	0.459	0.105	0.470	167	193
青色	0.368	0.155	0.646	0.664	283.5	76.5
绿色	0.308	0.304	0.541	0.620	240.5	119.5

（续）

	亮度（V）	峰－峰色度			色度相位（度）	
		U色轴（V）	V色轴（V）	总计（V）	行n（135°色同步信号）	行（n+1）（225°色同步信号）
洋红	0.217	0.304	0.541	0.620	60.5	299.5
红色	0.157	0.155	0.646	0.664	103.5	256.5
蓝色	0.060	0.459	0.105	0.470	347	13.0
黑色	0	0	0	0	–	–

表8-15 100/0/100/0（100%）PAL彩条

	亮度（V）	峰－峰色度			色度相位（度）	
		U色轴（V）	V色轴（V）	总计（V）	行n（135°色同步信号）	行（n+1）（225°色同步信号）
白色	0.700	0	–	–	–	–
黄色	0.620	0.612	0.140	0.627	167	193
青色	0.491	0.206	0.861	0.885	283.5	76.5
绿色	0.411	0.405	0.721	0.827	240.5	119.5
洋红	0.289	0.405	0.721	0.827	60.5	299.5
红色	0.209	0.206	0.861	0.885	103.5	256.5
蓝色	0.080	0.612	0.140	0.627	347	13.0
黑色	0	0	0	0	–	–

表8-16 100/0/100/25（98%）PAL彩条

	亮度（V）	峰－峰色度			色度相位（度）	
		U色轴（V）	V色轴（V）	总计（V）	行n（135°色同步信号）	行（n+1）（225°色同步信号）
白色	0.700	0	–	–	–	–
黄色	0.640	0.459	0.105	0.470	167	193
青色	0.543	0.155	0.646	0.664	283.5	76.5
绿色	0.483	0.304	0.541	0.620	240.5	119.5
洋红	0.392	0.304	0.541	0.620	60.5	299.5
红色	0.332	0.155	0.646	0.664	103.5	256.5
蓝色	0.235	0.459	0.105	0.470	347	13.0
黑色	0	0	0	0	–	–

为便于参考，表8-17、表8-18和表8-19示范了生成标准PAL彩条的RGB和YCbCr值。假定RGB值的范围为0~255，假定YCbCr值的范围是Y为16~235，Cb和Cr为16~240。

表8-17 75% PAL彩条的RGB和YCbCr值

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
gamma校正的RGB（gamma＝1/0.45）								
R'	255	191	0	0	191	191	0	0
G'	255	191	191	191	0	0	0	0
B'	255	0	191	0	191	0	191	0
线性RGB								
R	255	135	0	0	135	135	0	0

（续）

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
G	255	135	135	135	0	0	0	0
B	255	0	135	0	135	0	135	0
				YCbCr				
Y	235	162	131	112	84	65	35	16
Cb	128	44	156	72	184	100	212	128
Cr	128	142	44	58	198	212	114	128

表8-18　100% PAL彩条的RGB和YCbCr值

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
			gamma校正的RGB（gamma = 1/0.45）					
R′	255	255	0	0	255	255	0	0
G′	255	255	255	255	0	0	0	0
B′	255	0	255	0	255	0	255	0
				线性RGB				
R	255	255	0	0	255	255	0	0
G	255	255	255	255	0	0	0	0
B	255	0	255	0	255	0	255	0
				YCbCr				
Y	235	210	170	145	106	81	41	16
Cb	128	16	166	54	202	90	240	128
Cr	128	146	16	34	222	240	110	128

表8-19　98% PAL彩条的RGB和YCbCr值

	白色	黄色	青色	绿色	洋红	红色	蓝色	黑色
			gamma校正的RGB（gamma = 1/0.45）					
R′	255	255	44	44	255	255	44	44
G′	255	255	255	255	44	44	44	44
B′	255	44	255	44	255	44	255	44
				线性RGB				
R	255	255	5	5	255	255	5	5
G	255	255	255	255	5	5	5	5
B	255	5	255	5	255	5	255	5
				YCbCr				
Y	235	216	186	167	139	120	90	16
Cb	128	44	156	72	184	100	212	128
Cr	128	142	44	58	198	212	114	128

8.4.2　EIA彩条（NTSC）

EIA彩条（图8-28和表8-10）属于EIA-189-A标准。7个彩条（灰色、黄色、青色、绿色、洋红、红色和蓝色）的幅度为75%、饱和度为100%。每个彩条的宽度都是扫描行有效部分的1/7。注意图8-28中和表8-10的黑条都不属于标准，这里示范出来只是为了参考。该彩条测试信号允许检查色调和色饱和度的准确度。

314
↓
317

8.4.3 EBU彩条（PAL）

EBU彩条与EIA彩条相似，只是EBU彩条使用了一个100 IRE的白色电平（参见图8-29和表8-14）。EBU彩条有6个彩条（黄色、青色、绿色、洋红、红色和蓝色）的幅度为75%、饱和度为100%，而白条的幅度为100%。每个彩条的宽度都是扫描行有效部分的1/7。注意图8-29和表8-14中的黑条都不属于标准，这里的示范只是为了参考。该彩条测试信号允许检查色调和色饱和度的准确度。

8.4.4 SMPTE条（NTSC）

这个对场进行分隔的测试信号构成如下：EIA彩条构成场的前2/3，反转蓝条（reverse blue bar）构成接下来的1/12场，PLUGE测试信号构成场的剩余部分。

8.4.5 反转蓝条

反转蓝条由EIA/EBU彩条的蓝色、洋红和青色彩条构成，但是以不同的顺序排列——蓝色、黑色、洋红、黑色、青色、黑色和白色。每个彩条的宽度都是扫描行有效部分的1/7。一般情况下，反转蓝条通常和EIA或EBU彩条一起在对场进行分隔的情形中使用，EIA/EBU彩条构成场的前3/4，反转蓝条构成场的剩余部分。这种按场进行分隔的安排可以简化彩色监视器对色度和色调的调整。

318 ~ 322

8.4.6 PLUGE

PLUGE（Picture Line-Up Generating Equipment，图像线性生成装置）信号是一种可视黑色基准信号，该测试信号有一个比黑色更黑的区域、一个黑色区域和一个没有黑色那么黑的区域。在监视器调整过程中，要对监视器的亮度进行调整，从而使黑色区域和比黑色更黑的区域彼此之间难以区别，没有黑色那么黑的区域稍微淡一点（对比度正常设置）。在测试过程中，通常添加额外的测试信号，例如白脉冲和调制IQ信号，从而方便测试和监视器的调整。

NTSC制式的PLUGE测试信号（如图8-31所示）由一个7.5 IRE（黑色电平）的基底电平与一个40 IRE的“−I”相位调制、一个100 IRE的白脉冲、一个40 IRE的“+Q”相位调制和3.5 IRE、7.5 IRE及11.5 IRE的基底电平组成。通常情况下，PLUGE信号用作SMPTE条的一部分。

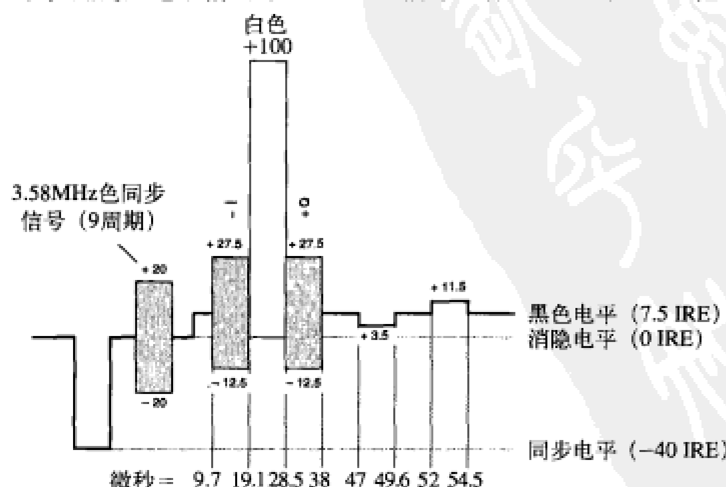


图8-31 NTSC制式的PLUGE测试信号。图中示范了IRE值

至于PAL制式，每个国家都有自己稍微有所不同的PLUGE信号配置，这些PLUGE信号的主要区别在于使用的黑色基底电平不同，并且每个国家都制定了自己的PLUGE标准测试信号。图8-32描述了一个典型的PAL制式PLUGE测试信号。PLUGE信号通常用作满屏测试信号，它由一个0 IRE的基底与PLUGE（－2 IRE、0 IRE和2 IRE的基底电平）和一个白色脉冲构成。白色脉冲根据扫描行数的不同，拥有5种电平亮度（0、25、50、75和100 IRE），如图8-32所示。PLUGE信号显示在拥有非0 IRE白色脉冲的扫描行上。ITU-R BT.1221中有关于各种PAL制式注意事项的讨论。

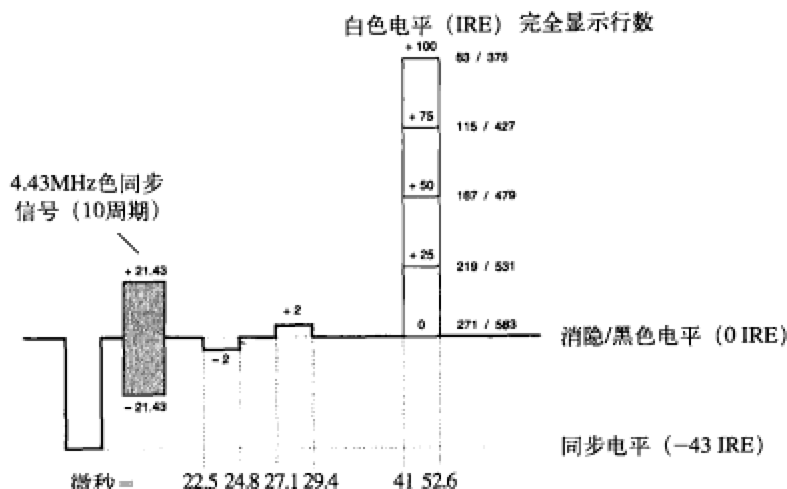


图8-32　PAL制式的PLUGE测试信号。图中示范了IRE值

8.4.7　Y条

Y条由仅含亮度电平的EIA/EBU彩条构成；但是，Y条包含有黑色电平（NTSC制式为7.5 IRE，PAL制式为0 IRE），并且仍然提供色同步信号。因此，每个亮度条的宽度是扫描行有效部分的1/8。Y条在彩色监视器调整和测试亮度非线性方面非常有用。一般情况下，Y条信号和EIA或EBU彩条信号一起在对场进行分隔的布局中使用，EIA/EBU彩条构成场的前3/4部分，Y条信号构成场的剩余部分。

8.4.8　红场

红场（red field）信号包含一个75%幅度、100饱和度的红色色度信号。这个测试信号非常有用，因为人的眼睛对混杂在红场中间的静态干扰信号非常敏感。导致图像质量中微小差错的变形可以通过视觉对图像的观察检查出来。一般情况下，红场信号和EIA/EBU彩条信号一起在对场进行分隔的布局中使用，EIA/EBU彩条构成场的前3/4部分，红场信号构成场的剩余部分。

8.4.9　10级阶梯

这个测试信号由10个未调制亮度构成，亮度之间间隔10 IRE，范围为0 IRE~100 IRE。如图8-33所示。该测试信号可以用于测试亮度非线性。

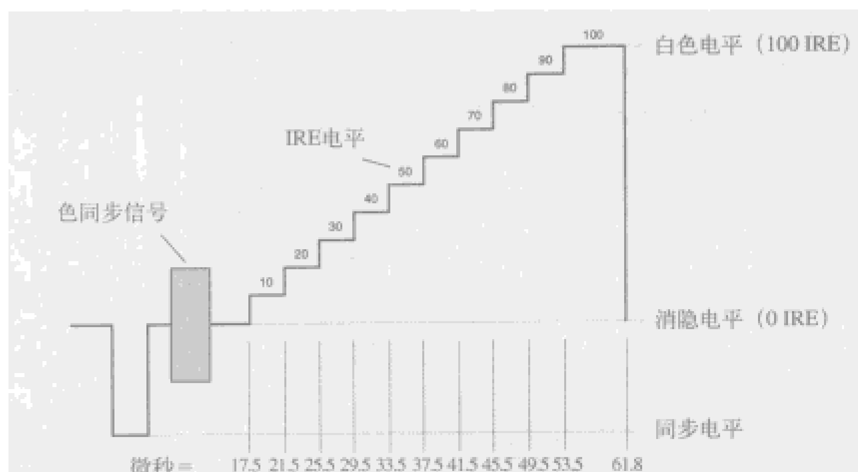


图8-33 NTSC和PAL制式的10级阶梯测试信号

8.4.10 调制斜坡

调制斜坡信号由一个从0 IRE到80IRE或者100 IRE的亮度斜坡构成，叠加有相对与色同步信号相位为0°±1°的调制色度，如图8-34所示。80 IRE斜坡可以用于测试系统的标准运算；而100 IRE斜坡可以自由地用于测试全部运算。（M）NTSC制式的峰－峰调制色度为40±0.5 IRE，（B、D、G、H、I）PAL制式的峰－峰调制色度为42.86±0.5 IRE。注意该测试信号中使用了一个0 IRE的台阶电平。调制斜坡包络起始点和终止点处的上升和下降时间为400±25ns（NTSC制式）或者近似为1μs（PAL制式）。该测试信号可以用于测试微分增益。当测试数字系统时，调制斜坡信号优于5级或10级调制阶梯信号。 325

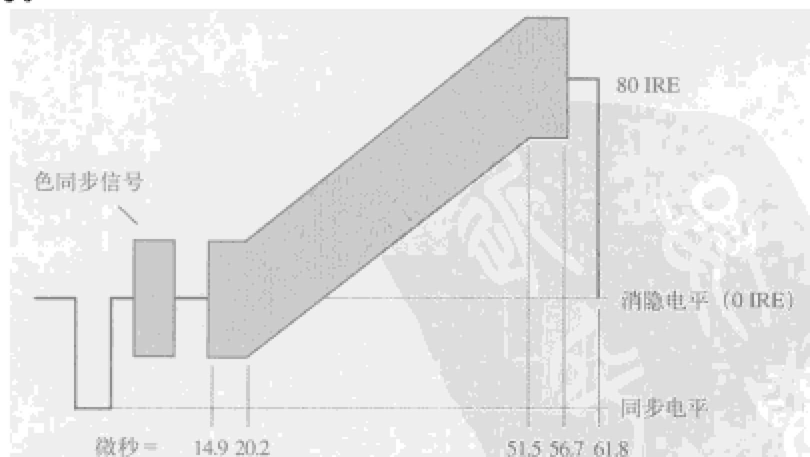


图8-34 NTSC和PAL制式的80 IRE调制斜坡测试信号

8.4.11 调制阶梯

5级调制阶梯信号（也有10级的），如图8-35所示，该信号包含5级亮度，叠加有相对于色同步信号相位为0°±1°的调制色度。（M）NTSC制式的峰－峰调制色度幅度为40±0.5 IRE，（B、D、G、H、I）PAL制式的峰–峰调制色度幅度为42.86±0.5 IRE。注意该测试信号中使用了一个0 IRE

的台阶电平。每个调制包包络的上升和下降时间为400±25ns（NTSC制式）或者近似为1μs（PAL制式）。5级调制阶梯信号的亮度IRE电平如图8-35所示。该测试信号可以用于测试微分增益。当测试数字系统时，调制斜坡信号优于5级或10级调制阶梯信号。

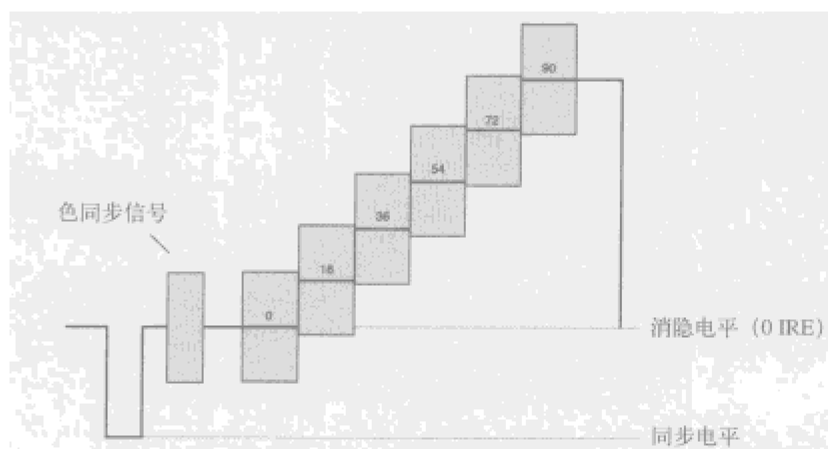
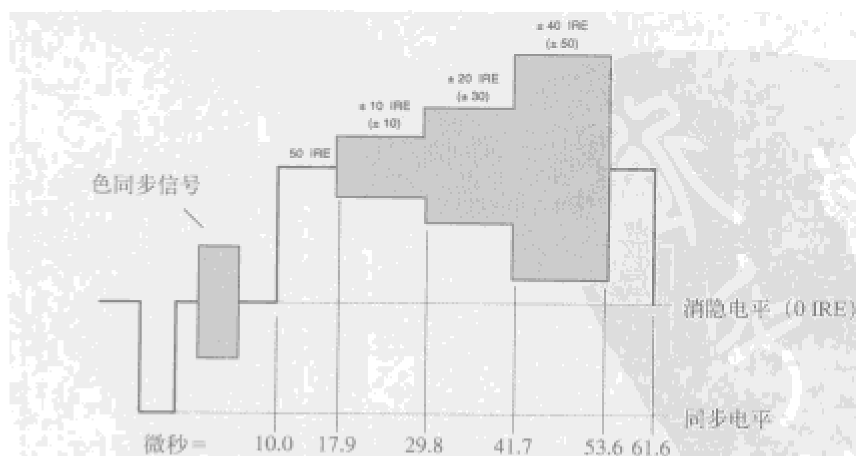


图8-35　NTSC和PAL制式的5级调制阶梯测试信号

8.4.12　调制基底

调制基底测试信号（也称为3级色度条），如图8-36所示，该信号包含一个50 IRE的亮度基底，叠加有3个相对于色同步信号相位为−90°±1°的调制色度的幅度。（M）NTSC制式的调制色度的峰−峰幅度为20±0.5、40±0.5和80±0.5 IRE，（B、D、G、H、I）PAL制式的调制色度的峰−峰幅度为20±0.5、60±0.5和100±0.5 IRE。注意该测试信号中使用了一个0 IRE的台阶电平。每个调制包包络的上升和下降时间为400±25ns（NTSC制式）或者近似为1μs（PAL制式）。该测试信号可以用于测试色度−亮度互调和色度非线性增益。

326



327

图8-36　NTSC和PAL制式的调制基底测试信号。括号中示范的是PAL制式的IRE值

8.4.13　多波群

（M）NTSC制式的多波群（multiburst）测试信号如图8-37所示，该信号由一个峰值幅度为100±1 IRE的白色标志和6个频率包构成，每个包拥有一个特定的频率。包拥有一个峰−峰幅度为60

±0.5 IRE的40±1 IRE的基底。注意该测试信号中使用了一个0 IRE的台阶电平，每个包的起始点和终止点位于0相位。

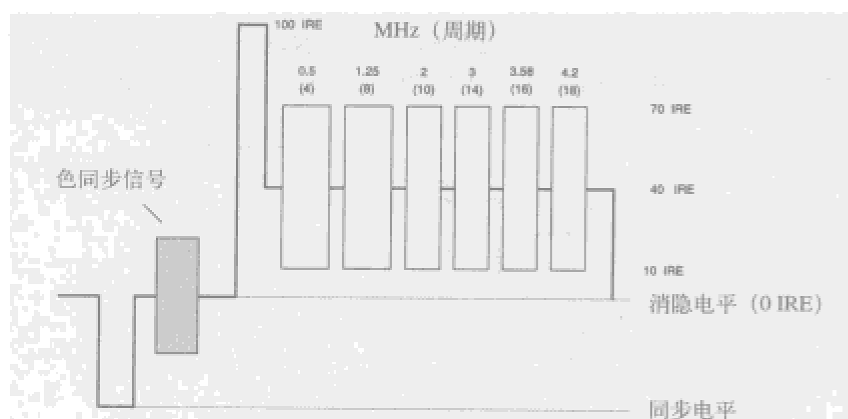


图8-37 NTSC制式的多波群测试信号

（B、D、G、H、I）PAL制式的ITU多波群测试信号如图8-38所示，该信号由一个峰值幅度为80±1 IRE的4μs白色标志和6个频率包构成，每个包拥有一个特定的频率。包拥有一个峰－峰幅度为60±0.5 IRE的50±1 IRE基底。注意该测试信号中每个包的起始点和终止点位于0相位。包之间的间隙为0.4~2.0μs。ITU多波群测试信号可以在第18行上提供。 328

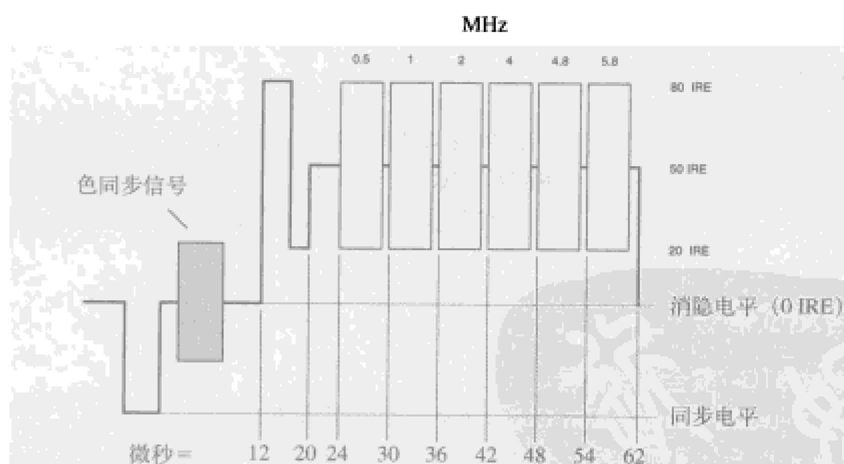


图8-38 PAL制式的ITU多波群测试信号

多波群信号通过测量包的峰－峰幅度来测试系统的频率响应。

8.4.14 线条

线条（line bar）是一个发生在有效扫描行时间之内（上升和下降时间都≤1μs）任何位置的10μs（PAL制式）、18μs（NTSC制式）或者25μs（PAL制式）单独100±0.5 IRE（基准白色）脉冲。注意该信号没有提供色同步信号，使用了一个0 IRE的台阶电平。该测试信号用来测量行时间失真（或者行倾斜）。数字编码器或者解码器不会生成行时间失真，行时间失真主要由模拟滤波器和传输通道生成。

8.4.15 多脉冲

（M）NTSC制式多脉冲测试信号包含一个2T脉冲和25T脉冲及12.5T脉冲加不同高频分量，如图8-39所示。（B、D、G、H、I）PAL制式多脉冲与（M）NTSC制式多脉冲相似，只是使用了20T和10T脉冲，而且没有7.5 IRE的台阶电平。该测试信号通常用来测量传输通道的频率响应。

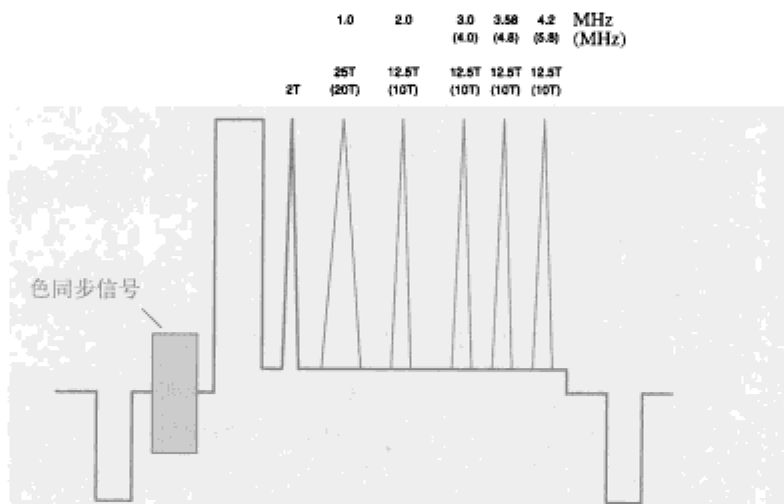


图8-39 NTSC和PAL制式的多脉冲测试信号。括号中示范的是PAL值

8.4.16 场方波

场方波测试信号包含场1全部有效行时间和场2消隐扫描行的100±0.5 IRE脉冲。注意场方波信号中没有提供色同步信号，并且使用的是0 IRE的台阶电平。该测试信号用于测量场时间失真（或者场倾斜）。数字编码器或者解码器不会产生场时间失真；场时间失真主要是由模拟滤波器和传输通道产生的。

8.4.17 复合测试信号

1. NTSC制式NTC-7复合测试信号

NTC（U.S. Network Transmission Committee，美国网络传输委员会）开发了一种复合测试信号，可以用来测试若干个视频参数，不必再使用多个测试信号。NTSC制式NTC-7复合测试信号（如图8-40所示）包含一个100 IRE的线条、一个2T脉冲、一个12.5T色度脉冲和一个5级调制阶梯信号。

329 ↓ 330

线条测试信号的峰值幅度为100±0.5 IRE，10%~90%上升和下降时间为125±5ns，形状为一个完整的正弦平方曲线。在60 IRE电平处的宽度为18μs。

2T脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为250±10ns。

12.5T色度脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为1562.5±50ns。

5级调制阶梯信号包含5级亮度，亮度上叠加有相对于色同步信号相位为0°±1°的40±0.5 IRE的副载波。每个调制包包络的上升和下降时间为400±25ns。

NTC-7复合测试信号可以在第17行上提供。

2. PAL制式ITU复合测试信号

ITU（BT.628和BT.473）开发了一种复合测试信号，可以用来测试若干个视频参数，不必再

使用多个测试信号。PAL制式的ITU复合测试信号（如图8-41所示）包含一个白色标志、一个2T脉冲和一个5级调制阶梯信号。

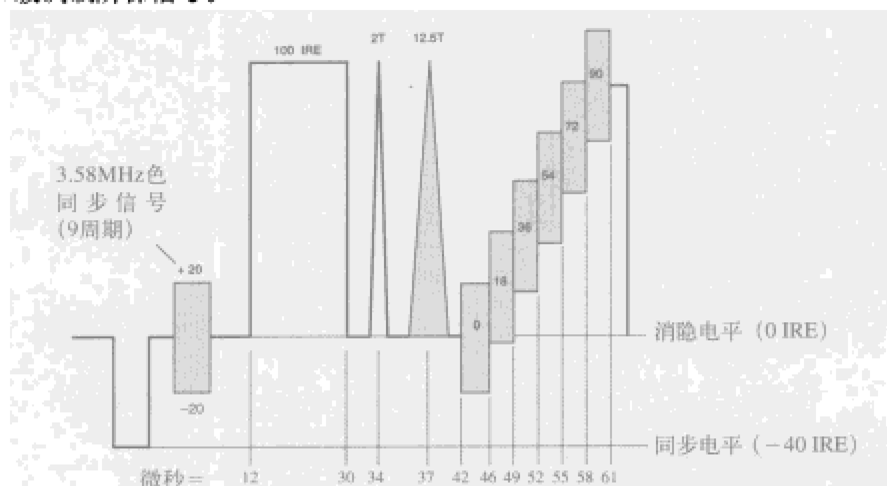


图8-40 NTSC制式NTC-7复合测试信号，图中示范了相应的IRE值

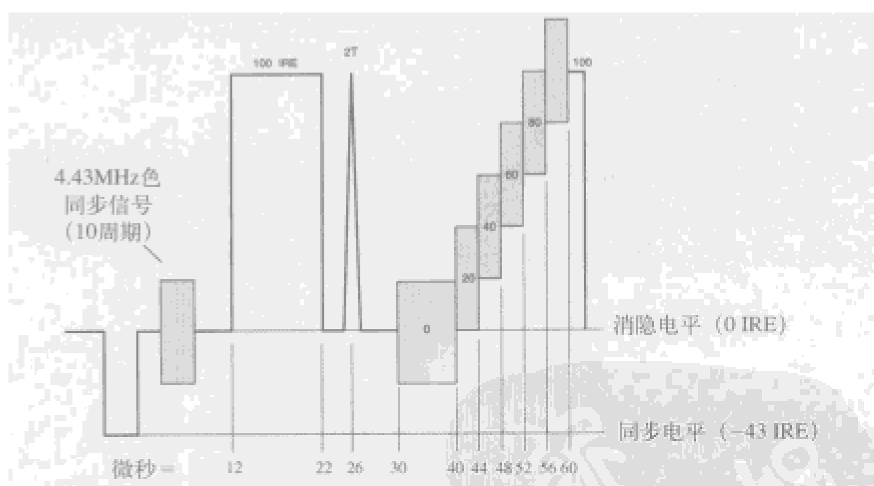


图8-41 PAL制式ITU复合测试信号，图中示范了相应的IRE值

白色标志的峰值幅度为100±1 IRE，宽度为10μs。

2T脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为200±10ns。

5级调制阶梯信号包含5级亮度，亮度上叠加有相对于U色轴相位为60°±1°的42.86±0.5 IRE的副载波（图8-41中示范了相应的IRE值）。每个调制包包络的上升和下降时间近似为1μs。 331

ITU复合测试信号可以在第330行上提供。

❑ 英国复合测试信号

因为10T脉冲要比20T脉冲对延迟误差更加敏感（以牺牲色度带宽为代价），所以英国允许采用一个稍微有所不同的测试信号。合适脉冲宽度的选择是尽可能完全占用PAL制色度带宽和脉冲对延迟误差尽可能敏感之间的折衷。因此，图8-42中的英国国家测试信号（由英国广播公司和英国独立电视管理委员会开发）可以在（I）PAL制式的第19行和332行上提供。

白色标志的峰值幅度为100±1 IRE，宽度为10μs。

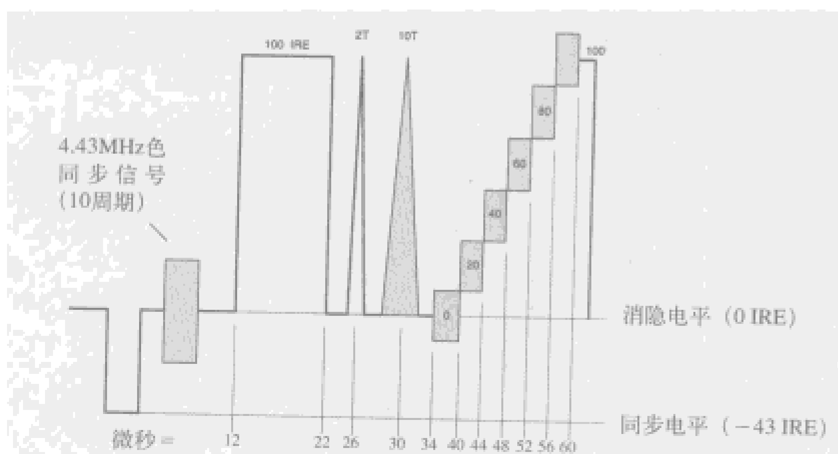


图8-42　英国（I）PAL制式国家测试信号#1，图中示范了相应的IRE值

2T脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为200±10ns。

10T色度脉冲的峰值幅度为100±0.5 IRE。

5级调制阶梯信号包含5级亮度，亮度上叠加有相对于U色轴相位为60°±1°的21.43±0.5 IRE的副载波（图8-42中示范了相应的IRE值）。每个调制包包络的上升和下降时间近似为1μs。

332

8.4.18　组合测试信号

1. NTSC制式NTC-7组合测试信号

NTC（U.S. Network Transmission Committee，美国网络传输委员会）还开发了一种组合测试信号，可以用来测试若干个视频参数，不必再使用多个测试信号。NTSC制式NTC-7组合测试信号（如图8-43所示）包含一个白色标志、一个多波群和一个调制基底信号。

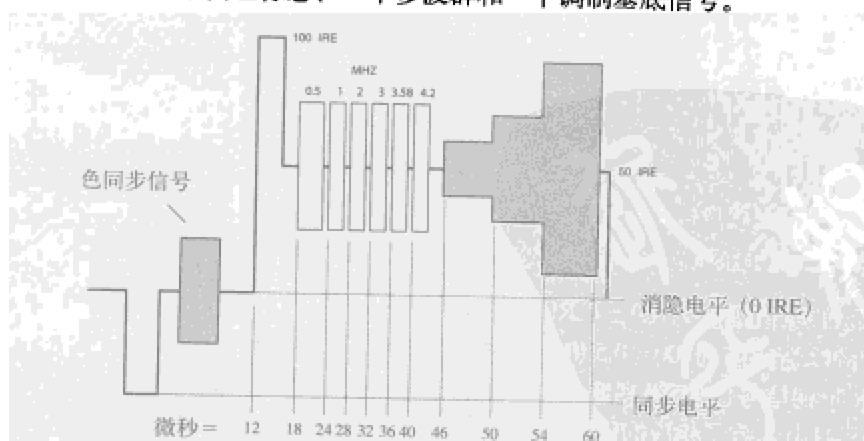


图8-43　NTSC制式的NTC-7组合测试信号

白色标志的峰值幅度为100±1 IRE，宽度为4μs。

多波群的基底为50±1 IRE，峰－峰幅度为50±0.5 IRE。每个频率包的起始点位于0相位。0.5MHz包的宽度为5μs，剩余包的宽度为3μs。

3级调制基底信号包含一个50 IRE亮度基底，叠加有3个相对于色同步信号相位为−90°±1°的调制色度的幅度（峰－峰幅度分别为20±0.5、40±0.5和80±0.5 IRE）。每个调制包包络的上升和

下降时间为400±25ns。

NTC-7组合测试信号可以在第280行上提供。

2. PAL制式ITU组合测试信号

ITU（BT.473）开发了一种组合测试信号，可以用来测试若干个视频参数，不必再使用多个测试信号。PAL制式ITU组合测试信号（如图8-44所示）包含一个白色标志、一个2T脉冲、一个20T调制色度脉冲和一个5级亮度阶梯信号。 [333]

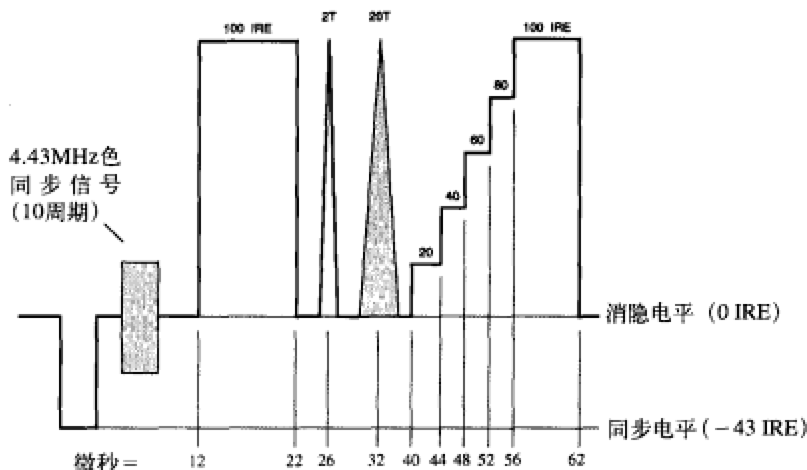


图8-44 PAL制式的ITU组合测试信号

线条信号的峰值幅度为100±1 IRE，宽度为10μs。

2T脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为200±10ns。

20T色度脉冲的峰值幅度为100±0.5 IRE，半幅度宽度为2.0±0.06μs。

5级亮度阶梯信号包含5级分别为20、40、60、80和100±0.5 IRE的亮度。

ITU组合测试信号可以在第17行上提供。

3. PAL制式ITU组合ITS测试信号

ITU（BT.473）开发了一种组合ITS（Insertion Test Signal，插入测试信号）测试信号，可以用来测试若干个PAL制式视频参数，不必再使用多个测试信号。PAL制式ITU组合ITS测试信号（如图8-45所示）包含一个峰－峰幅度分别为20、60和100±1 IRE的3级调制基底和一个峰－峰幅度为60±1 IRE的扩展副载波包。每个副载波包包络的上升和下降时间近似为1μs。每个副载波包的相位相对于U色轴为60°±1°。50 IRE电平上的容限为±1 IRE。

ITU复合ITS测试信号可以在第331行提供。

❑ 英国组合测试信号

英国允许使用稍微有所不同的测试信号，如图8-46所示。英国采用的（I）PAL制式在第20行和第333行上提供该组合测试信号。

该测试信号包含一个50 IRE亮度条，部分叠加有一个相对于U色轴相位为60°±1°的100 IRE副载波，以及一个位于第二半扫描行上的副载波的扩展色同步信号。 [334]

8.4.19 T脉冲

拥有快速上升时间的方波无法用于测试视频系统，因为带外分量的衰减和相位偏移致使输出信号中存在嗡嗡的响声，使要测量的带内失真模糊。T或者正弦平方脉冲的带宽受限，所以可以

用于测试视频系统。

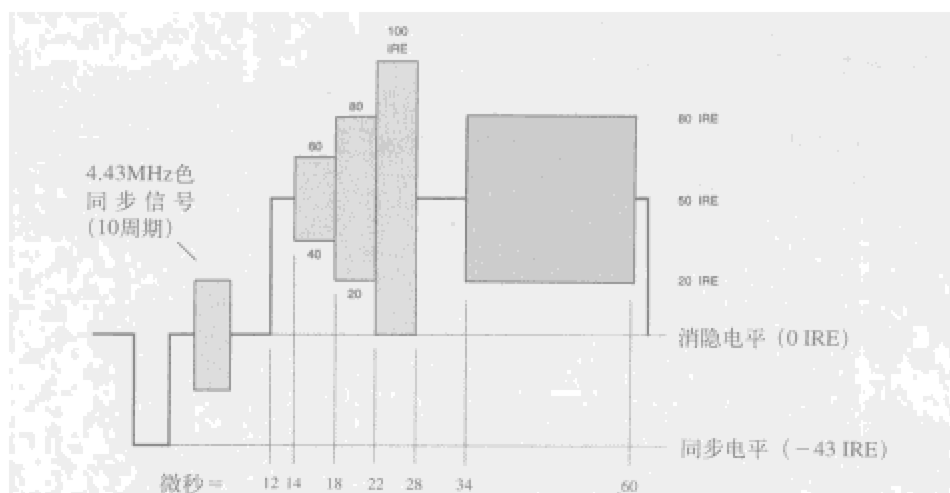
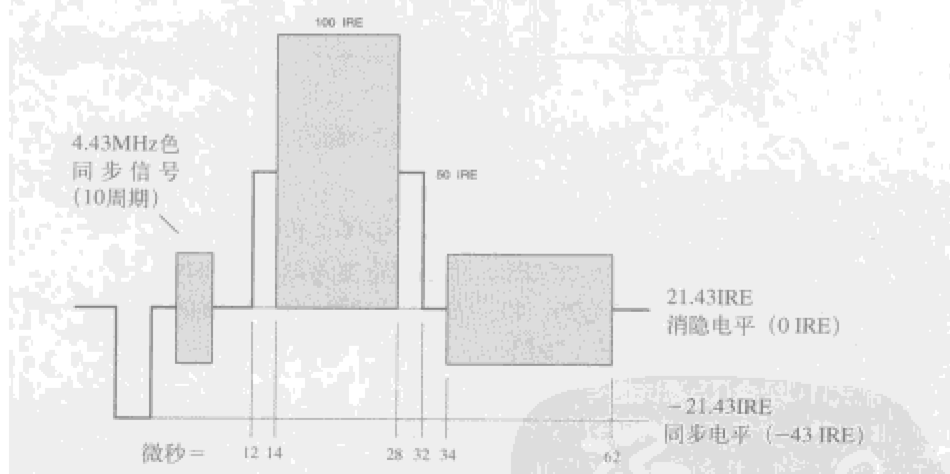


图8-45　PAL制式ITU组合ITS测试信号



335

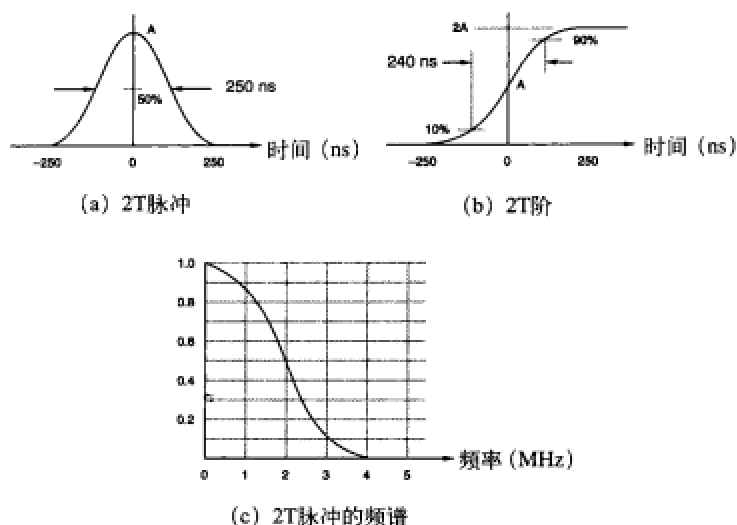
图8-46　英国（I）PAL制式国家测试信号#2

2T脉冲如图8-47所示，它就像T脉冲一样，经半周期余弦波平方计算而得。T脉冲根据半幅宽度（Half Amplitude Duration，HAD）指定，HAD是脉冲幅度50%处的测量脉冲宽度。拥有HAD的脉冲是时间间隔T的倍数，可以用来测试视频系统。如图8-39至图8-44所示，当测试NTSC制式视频系统时，通常使用的是T、2T、12.5T和25T脉冲，而PAL制式视频系统通常使用的是T、2T、10T和20T脉冲。

T是奈奎斯特间隔（Nyquist interval）或者

$1/2F_C$

其中F_C是视频系统的截止频率。对于NTSC制式来说，F_C是4MHz，而对于PAL制式来说，F_C是5MHz。因此，NTSC制式的T是125ns，而PAL制式的T是100ns。对于拥有一个125ns的HAD的T脉冲来说，2T脉冲的HAD为250ns，依此类推。2T脉冲的频谱如图8-47所示，代表典型字符发生器波形的能量含量。



（a）2T脉冲

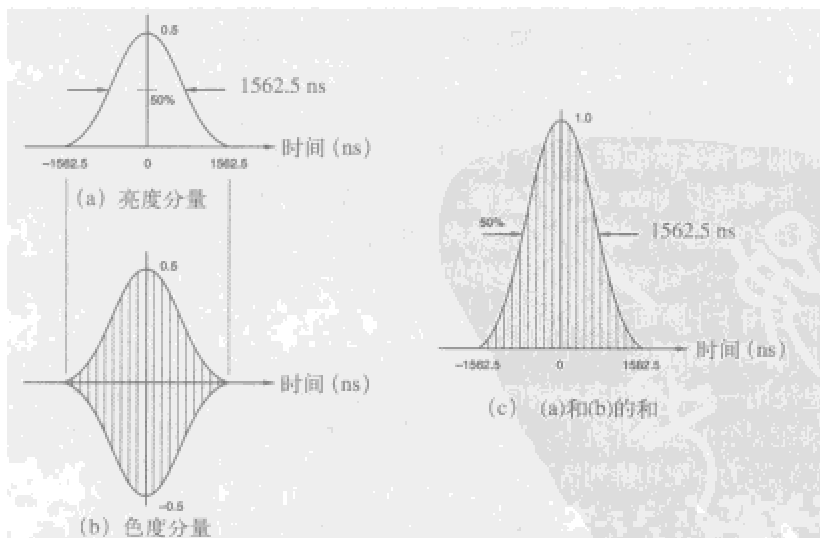
（b）2T阶

（c）2T脉冲的频谱

图8-47 T脉冲

为了生成大多数视频信号平滑的上升和下降边，通常使用T阶（通过集成T脉冲生成）信号。T阶信号的10%~90%上升/下降时间为0.964T，并且拥有一个界限分明的带宽。2T阶信号由图8-47中所示的2T脉冲生成。 336

12.5T色度脉冲（如图8-48所示）是一个很好的测试信号，可以用来测试任何色度－亮度时序误差，因为它的能量频谱分布聚集在两个相对较窄的带上。通过使用这个信号，可以检测亮度和色度相位失真的差别，但是不能检测其他频率组。



（a）亮度分量

（b）色度分量

（c）（a)和(b)的和

图8-48 12.5T色度脉冲

8.5 VBI数据

VBI（Vertical Blanking Interval，场消隐期，也叫场逆程）数据可以插入到有效图像区域的最高大约5个扫描行上，从而确保它不会被替换VBI的设备、删除VBI的DSS MPEG或者插入自己

VBI数据的有线电视系统删除。这是尼尔逊（Neilson）和其他确保它们的节目和商业跟踪数据通过布线系统到达接收机的常见做法。在大多数情况下，VBI数据是不可见的，因为它被电视机的过扫描遮蔽。

8.5.1 时间码

较常使用的时间编码有两种，由ANSI/SMPTE 12M和IEC 461定义：纵向时间码（Longitudinal 337 Time Code，LTC）和垂直消隐期时间码（Vertical Interval Time Code，VITC）。

LTC在单独的音轨上记录；因此，模拟VCR必须使用高带宽的放大器和音频磁头。这是因为时间码频率随着磁带速度的增加而增加，直至系统的频率响应导致产生无法可靠读取的失真时间码信号。随着磁带速度的降低，时间码频率也降低，直至在特别慢的磁盘速度或者停止图像情况下时间码信息再也无法获取到。

VITC作为视频信号的一部分记录；因此，无论磁带的速度是多少，时间码信息总是可用。但是，LTC允许在不写入视频信号的情况下就可以写入时间码信号；如果时间码信息需要变化，VITC要求视频信号也要变化。因此LTC对于同步多个音频或者音/视频源非常有用。

1. 丢帧

如果场频是每秒60/1.001场，那么直接按每秒60场统计就会产生这样一种错误，即每运行一小时要多统计108帧。这种问题的处理方法有以下3种。

非丢帧（nondrop frame）：在一个连续记录期间，每次统计增加1帧。在这种模式中，丢帧标志为0。

丢帧（drop frame）：为了最小化时序误差，除了第00、10、20、30、40和50分钟，其他分钟处每分钟起始点处的前两个帧的编号（00和01）从统计中忽略。在这种模式中，丢帧标志为1。

（M）PAL的丢帧：为了最小化时序误差，除了第00、20和40分钟，其他分钟处每两分钟（偶数分钟编号）每分钟起始处的前4个帧的编号（00~03）从统计中忽略。在这种模式中，丢帧标志为1。

即便采用了丢帧处理技术，在长期运行情况下，每24小时大约还有2.26帧的长期误差。如果它们与正确的时间有关联的话，这个误差累积必须定期复位。一般情况下，"复位到真实时间"指的是一个"阻塞同步（jam sync）"过程。一些阻塞同步实现将时间码复位为00：00：00.00，因此，复位必须在午夜进行。其他方法允许真正的同步，同步到正确的时间。

阻塞同步校正一个固有的问题就是时间码的中断。尽管该不连续性可能极短，但是它可能导致时间码读取器由于中断而"打嗝"。

2. 纵向时间码（LTC）

LTC信息使用一条单独的串行接口传输，和AES/EBU数字音频接口标准使用相同的电子接口，LTC信息记录在一个单独的音轨上。LTC信息时间数据的基本结构基于BCD系统。表8-20和表8-21列举了LTC信息位的分配和排列情况。注意这里使用的是24小时制时钟。

表8-20 LTC信息位的分配

位	功 能	注 释	位	功 能	注 释
0~3	帧的个位		58	标志5	注5
4~7	用户组1		59	标志6	注6
8~9	帧的十位		60~63	用户组8	
10	标志1	注1	64	同步位	固定"0"
11	标志2	注2	65	同步位	固定"0"

（续）

位	功能	注释	位	功能	注释
12~15	用户组2		66	同步位	固定“1”
16~19	秒的个位		67	同步位	固定“1”
20~23	用户组3		68	同步位	固定“1”
24~26	秒的十位		69	同步位	固定“1”
27	标志3	注3	70	同步位	固定“1”
28~31	用户组4		71	同步位	固定“1”
32~35	分钟的个位		72	同步位	固定“1”
36~39	用户组5		73	同步位	固定“1”
40~42	分钟的十位		74	同步位	固定“1”
43	标志4	注4	75	同步位	固定“1”
44~47	用户组6		76	同步位	固定“1”
48~51	小时的个位		77	同步位	固定“1”
52~55	用户组7		78	同步位	固定“0”
56~57	小时的十位		79	同步位	固定“1”

注：
1. 丢帧标志。525行和1125行系统：如果存在丢帧编号，为1；如果没有丢帧，为0。625行系统：0。
2. 彩色帧标志。525行系统：如果偶帧个位编号标记场1和场2，奇帧个位编号标记场3和场4，为1。625行系统：如果时间码根据8场序列锁定到视频信号并且视频信号拥有“优选行同步副载波相位（subcarrier-to-line-sync phase）”，为1。1125行系统：0。
3. 525行和1125行系统：相位校正。这个位应该设置一个状态，从而每80比特的字就包含一个偶数0。625行系统：二进制组标志0。
4. 525行和1125行系统：二进制组标志0。625行系统：二进制组标志2。
5. 二进制组标志1。
6. 525行和1125行系统：二进制组标志2。625行系统：相位校正。这个位应该设置一个状态，从而每80bit的字就包含一个偶数0。

表8-21 LTC信息位的排列

帧（525行和1125行系统计数0~29，625行系统计数0~24）	
帧的个位（位0~3） 帧的十位（位8~9）	4位BCD（计数0~9），位0是LSB 2位BCD（计数0~2），位8是LSB
秒	
秒的个位（位16~19） 秒的十位（位24~26）	4位BCD（计数0~9），位16是LSB 3位BCD（计数0~5），位24是LSB
分钟	
分钟的个位（位32~35） 分钟的十位（位40~42）	4位BCD（计数0~9），位32是LSB 3位BCD（计数0~5），位40是LSB
小时	
小时的个位（位48~51） 小时的十位（位56~57）	4位BCD（计数0~9），位48是LSB 2位BCD（计数0~2），位56是LSB

❑ LTC时序

这里采用以下调制技术：传输发生在每个比特周期（bit period）的起始点。当从比特起始点的半比特周期处开始第二次传输时表示1。当比特周期内没有传输时表示0。（参见图8-49）。信号的峰－峰幅度为0.5~4.5V，上升和下降时间为40±10μs（10%~90%幅度点）。

因为整个帧时间用于生成80bit的LTC信息，所以比特率（bit/s）可以由下式计算：

338

$F_C = 80F_V$

其中F_V是垂直场帧频，单位是帧/秒。80比特的时间码信息串行输出，第0比特首先输出。LTC字占用整个帧时间，数据必须在这个时间内平均分布。对于525行系统，LTC字的起始点位于第5±1.5行的开头；对于625行系统，LTC字的起始点位于第2±1.5行的开头；而对于1125行系统，LTC字的起始点位于帧的垂直同步时序基准±1行处。

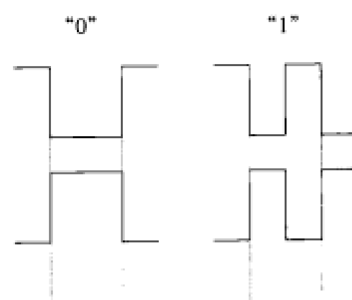


图8-49　LTC数据比特传输格式

3. 垂直消隐期时间码（VITC）

VITC在两场的视频信号垂直消隐期中记录。因为它和视频一起记录，所以可以在静止模式中读取。但是，它不能重录[或者重分条（restriped）]。重分条要求重新配音，删除并插入一个新的时间码。对于YPbPr和S-video接口，VITC在Y信号上提供。对于模拟RGB接口，VITC在全部3个信号上提供。

和LTC一样，VITC时间数据的基本结构也基于BCD系统。表8-22和表8-23列举了VITC信息位的分配和排列情况。注意这里使用的是24小时制时钟。

表8-22　VITC信息位的分配

位	功　能	注　释	位	功　能	注　释
0	同步位	固定"1"	42~45	分钟的个位	
1	同步位	固定"0"	46~49	用户组5	
2~5	帧的个位		50	同步位	固定"1"
6-9	用户组1		51	同步位	固定"0"
10	同步位	固定"1"	52~54	分钟的十位	
11	同步位	固定"0"	55	标志4	注4
12~13	帧的十位		56~59	用户组6	
14	标志1	注1	60	同步位	固定"1"
15	标志2	注2	61	同步位	固定"0"
16~19	用户组2		62~65	小时的个位	
20	同步位	固定"1"	66~69	用户组7	
21	同步位	固定"0"	70	同步位	固定"1"
22~25	秒的个位		71	同步位	固定"0"
26~29	用户组3		72~73	小时的十位	
30	同步位	固定"1"	74	标志5	注5
31	同步位	固定"0"	75	标志6	注6
32~34	秒的十位		76~79	用户组8	
35	标志3	注3	80	同步位	固定"1"
36~39	用户组4		81	同步位	固定"0"
40	同步位	固定"1"	82~89	CRC组	
41	同步位	固定"0"			

注：
1. 丢帧标志。525行和1125行系统：如果存在丢帧编号，为1；如果没有丢帧，为0。625行系统：0。
2. 彩色帧标志。525行系统：如果偶帧个位编号标记场1和场2，奇场个位编号标记场3和场4，为1。625行系统：如果时间码根据8场序列锁定到视频信号并且视频信号拥有"优选行同步副载波相位"，为1。1125行系统：0。
3. 525行系统：场标志。场1和场3期间为0，场2和场4期间为1。625行系统：二进制组标志0。1125行系统：场标志。场1期间为0，场2期间为1。
4. 525行和1125行系统：二进制组标志0。625行系统：二进制组标志2。
5. 二进制组标志1。
6. 525行和1125行系统：二进制组标志2。625行系统：场标志。场1、3、5和7期间为0。场2、4、6和8期间为1。

表8-23 VITC信息位的排列

帧（525行和1125行系统计数0~29，625行系统计数0~24）	
帧的个位（位2~5）	4位BCD（计数0~9），位2是LSB
帧的十位（位12~13）	2位BCD（计数0~2），位12是LSB
秒	
秒的个位（位22~25）	4位BCD（计数0~9），位22是LSB
秒的十位（位32~34）	3位BCD（计数0~5），位32是LSB
分钟	
分钟的个位（位42~45）	4位BCD（计数0~9），位42是LSB
分钟的十位（位52~54）	3位BCD（计数0~5），位52是LSB
小时	
小时的个位（位62~65）	4位BCD（计数0~9），位62是LSB
小时的十位（位72~73）	2位BCD（计数0~2），位72是LSB

❑ VITC循环冗余检验

VITC信息中保留了8位（82~89）用于码字错误检查，错误检查通过循环冗余检验。生成多项式(x^8+1)，可以应用于所有位［从0到81（包括0和81）］。图8-50描述了使用一个移位寄存器的多项式的实现。在时间码数据通路中，多路复用器位于位置0，在移位寄存器计算CRC值时同步输出数据。在所有的时间码数据输出之后，移位寄存器就包含CRC值，然后将多路复用器切换到位置1，从而使CRC值输出。当在解码过程中重复这一过程时，如果没有错误发生，移位寄存器所包含的值应该都为0。

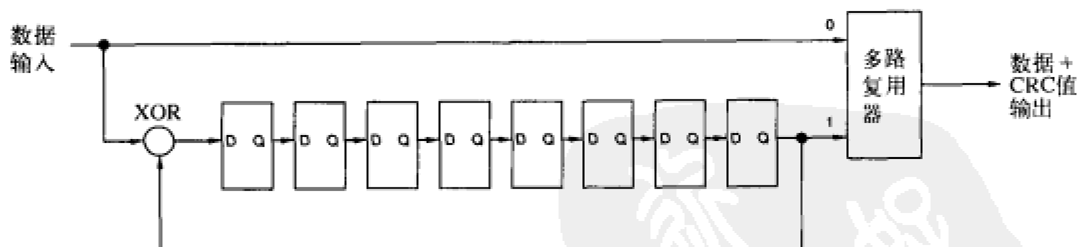


图8-50 VITC信息CRC值的生成

❑ VITC时序

这里采用的调制技术是每个状态对应一个二进制状态，仅当邻接位有从1到0或者从0到1的转换时，状态才会发生转变。当邻接位包含相同的数据时，没有转变发生。这通常称为“不归零制(non-return to zero，NRZ)”。同步比特对插入到VITC数据中，帮助接收机维护正确的频率锁定(frequency lock)。

比特率（F_C）定义如下：

$F_C=115F_H\pm2\%$

其中F_H是水平频。90bit的时间码信息串行输出，第0位首先输出。对于625i（576i）系统来说，第19行和第332行（或者第21行和第334行）通常用于VITC。对于525i（480i）系统来说，通常使用第14行和第277行。对于1125i（1080i）系统来说，通常使用第9行和第571行。尽管VITC信息可以使用每一场上的任意两个非连续的扫描行，但是为了防止VITC信息的漏失，还可以在后面的两

个扫描行上提供它。

图8-51描述了扫描行上VITC数据的时序。VITC数据必须平均地分布在VITC字中。VITC信息位的数据的10%~90%上升和下降时间在通过下游的色度电路添加到视频中之前应该为200±50ns（525i和625i系统），或者100±25ns（1125i系统），以避免VITC信号可能失真。在大多数环境中，视频D/A转换器之后的模拟低通滤波器应该有能力胜任滤波任务。

339~341

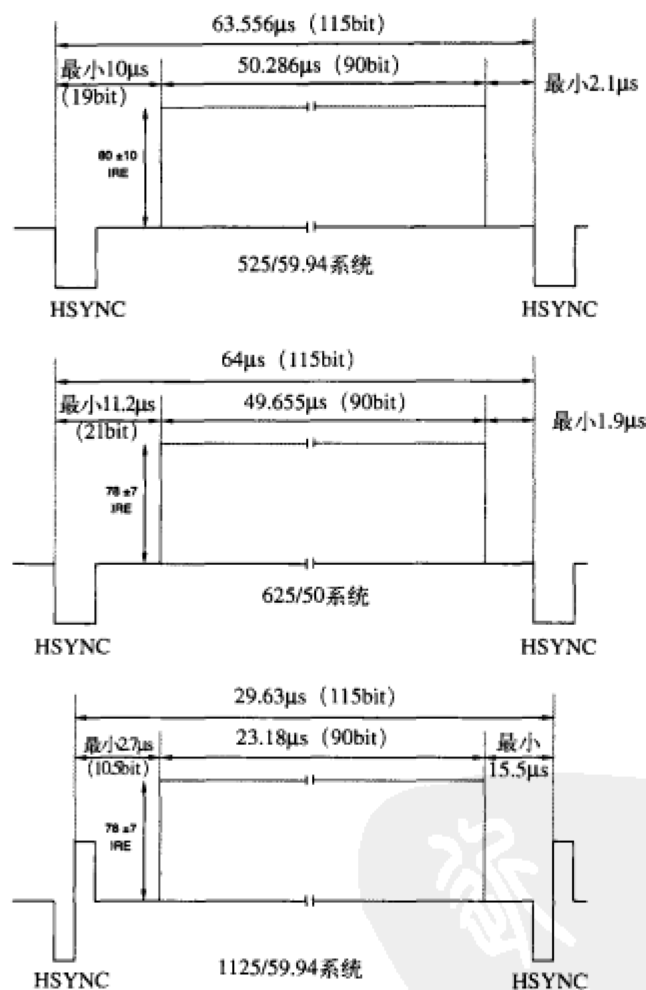


图8-51　VITC位置和时序

4. 用户位

二进制组标志（Binary Group Flag，BGF）的各位如表8-24所示，指定了32个用户位的内容。32个用户位分成8组，每组4 bit。

表8-24　LTC和VITC二进制组标志（BGF）位定义

用户位内容	外部时钟的时间码参考	BGF2	BGF1	BGF0
用户自定义	否	0	0	0
8位字符集[1]	否	0	0	1
用户自定义	是	0	1	0
保留	未分配	0	1	1

（续）

用户位内容	外部时钟的时间码参考	BGF2	BGF1	BGF0
日期和时区[3]	否	1	0	0
页/行[2]	否	1	0	1
日期和时区[3]	是	1	1	0
页/行[2]	是	1	1	1

注：1. 遵循ISO/IEC 646或者2022。
2. 在SMPTE 262M中描述。
3. 在SMPTE 309M中描述。参见表8-25至表8-27。

用户位的目的是为了用户存储数据。如果通过二进制组标志指定按用户自定义方式分配，那么32位可以按任意的方式分配，不受任何限制。

如果二进制组标志指定了一个遵循ISO/IEC 646或者2022的8位字符集，那么字符按照图8-52所示插入。注意一些用户位在二进制组标志解码之前解码，因此，解码器必须在任何处理完成之前存储早先的用户数据。

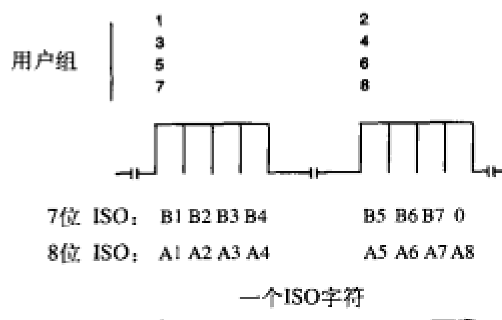


图8-52 使用二进制组描述7位或8位编码的ISO字符

当用户组用于传送时区和日期信息时，用户组7和8在剩余6个用户组中指定时区和日期的格式，如表8-25和表8-27所示。日期可以是6位数字YYMMDD的格式（表8-26），也可以是6位数字的改进儒略日①（Julian date）（MJD），这由MJD标志指定。

表8-25 日期和时区格式编码

用户组8				用户组7			
位3	位2	位1	位0	位3	位2	位1	位0
MJD标志	0	时区偏移码0x00-0x3F					

注：MJD标志：0＝YYMMDD格式，1＝MJD格式。

表8-26 YYMMDD日期格式

用 户 组	分 配	值	描 述
1	D	0~9	日的个位
2	D	0~3	日的十位

① 计算机系统中的日历表示法，采用年和一年中经过的天数表示，例如96年3月18日写成96.78。

（续）

用户组	分配	值	描述
3	M	0~9	月的个位
4	M	0、1	月的十位
5	Y	0~9	年的个位
6	Y	0~9	年的十位

表8-27　时区偏移编码

代码	小时	代码	小时	代码	小时
00	UTC	16	UTC+10.00	2C	UTC+09.30
01	UTC-01.00	17	UTC+09.00	2D	UTC+08.30
02	UTC-02.00	18	UTC+08.00	2E	UTC+07.30
03	UTC-03.00	19	UTC+07.00	2F	UTC+06.30
04	UTC-04.00	1A	UTC-06.30	30	TP-1
05	UTC-05.00	1B	UTC-07.30	31	TP-0
06	UTC-06.00	1C	UTC-08.30	32	UTC+12.45
07	UTC-07.00	1D	UTC-09.30	33	保留
08	UTC-08.00	1E	UTC-10.30	34	保留
09	UTC-09.00	1F	UTC-11.30	35	保留
0A	UTC-00.30	20	UTC+06.00	36	保留
0B	UTC-01.30	21	UTC+05.00	37	保留
0C	UTC-02.30	22	UTC+04.00	38	用户自定义
0D	UTC-03.30	23	UTC+03.00	39	未知
0E	UTC-04.30	24	UTC+02.00	3A	UTC+05.30
0F	UTC-05.30	25	UTC+01.00	3B	UTC+04.30
10	UTC-10.00	26	保留	3C	UTC+03.30
11	UTC-11.00	27	保留	3D	UTC+02.30
12	UTC-12.00	28	TP-3	3E	UTC+01.30
13	UTC+13.00	29	TP-2	3F	UTC+00.30
14	UTC+12.00	2A	UTC+11.30		
15	UTC+11.00	2B	UTC+10.30		

8.5.2　CEA-608隐藏字幕

本节回顾美国为听力损伤者提供的CEA-608隐藏字幕（closed captioning）。隐藏字幕和文本在第21行和第284行的消隐有效行时间期间传送。但是，由于视频编辑，它们偶尔可能位于第21~25行和第284~289行上的任何一行上。

342 ~ 346

扩展数据服务（extended data service，XDS）包也可以在第284行的消隐有效行时间期间传送。XDS包可以表示节目名称、放映时间、剩余时间等信息。

注意由于编辑在传输之前，所以字幕信息有可能偶尔会向下移一个或者两个扫描行。因此，为了获取字幕信息，字幕解码器应该监视的不止第21行和第284行。

1. 波形

两行的数据格式都包含一个时钟插补（run-in）信号、一个起始位和两个7位的ASCII数据加上ASCII数据的奇偶校验字（根据X3.4-1967）。对于YPbPr和S-video接口，字幕在Y信号上提供。对于模拟RGB接口，字幕在绿色通道上提供，幅度为复合信号或者Y信号的1.7倍。

图8-53描述了传送隐藏字幕和XDS信息的波形和时序信息，它遵循CEA-608。时钟插补信号是一个相对于字幕数据频率锁定和相位锁定的7周期的正弦色同步信号，用来为解码器提供同步信号。时钟插补信号的标称数据速率为32 × F_H。但是，解码器不应该依赖于这种时序关系，因为视频处理电路和VCR可能会引入水平时序的变化。在时钟插补信号之后，维护了两个数据位同期的消隐电平，后面跟着一个起始位1。起始位后面是16 bit的数据，包括两个7 bit的ASCII字符和两个ASCII字符的奇校验。字幕数据使用不归零制（Non-Return-to-Zero，NRZ）码传送，1对应于50 ± 2 IRE的电平，0对应于消隐电平（0~2 IRE）。时钟的反向交叉与数据位传送相关。

347~348

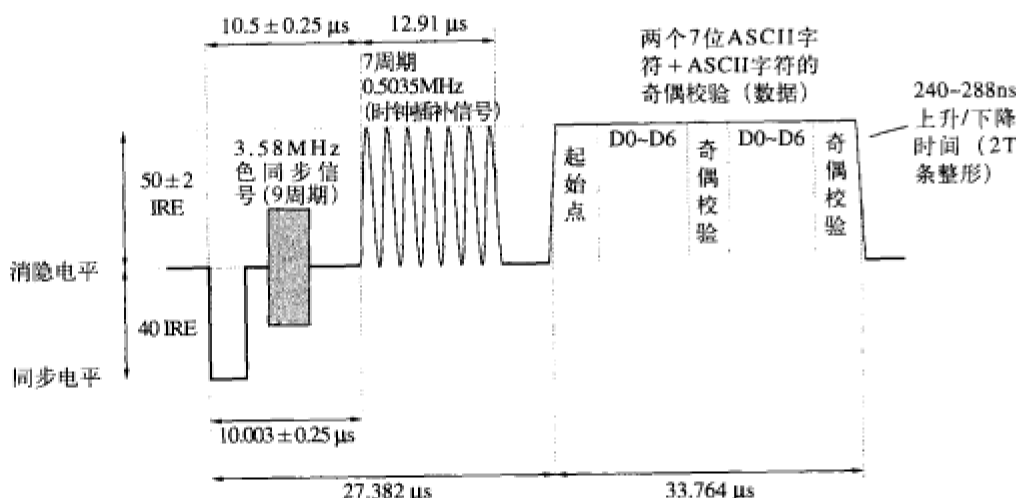


图8-53 525行系统第21行和第284行隐藏字幕时序

典型的解码器指定同步信号50%位置处和时钟插补信号之间的时间为10.5 ± 0.5μs，容限为F_H的 ± 3%，位1代表50 ± 12 IRE，位0代表 − 2至 + 12 IRE。解码器还必须处理240~480ns的位上升/下降时间。

当不传送显示或控制字符时，应该发送NUL字符（0x00）。这个字符和时钟插补信号结合在一起，允许解码器确定执行的是字幕传送还是文本传送。

如果只使用第21行，那么时钟插补信号和数据不需要在第284行上提供。但是，如果只使用第284行，那么时钟插补信号和数据应该在第21行和第284行上都提供；第21行上的数据将包含NUL字符。

在解码器上，如图8-54所示，525行4：3隔行显示的显示区域通常有15排高和34列宽。显示区域垂直方向从第43行和第306行开始，在第237行和第500行上结束。显示区域水平方向从13μs开始，到58μs结束，与水平同步信号的前沿（leading edge）一致。

在文本模式中，所有的排都用于显示文本；每排最多包含32个字符，文本的左边和右边至少要有一列宽度的空间。唯一的透明区域就是文本区域的外围。

在字幕模式中，文本通常只显示在第1~4排或者第12~15排上；剩余排通常是透明的。每排最多包含32个字符，文本的左边和右边至少要有一列宽度的空间。

一些字幕解码器支持每排多达48列，最多16排，允许对字幕数据的显示进行一些定制。

2. 基本服务

基本服务有两种类型：文本模式（一种与节目通常不相关的数据服务）和字幕。

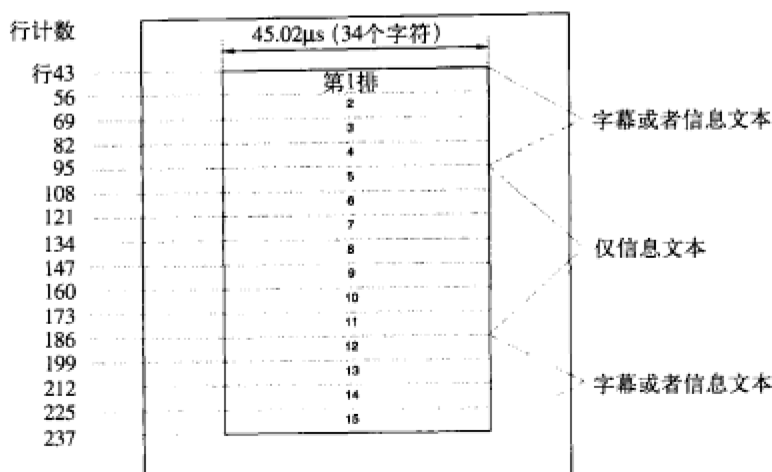


图8-54 隐藏字幕显示格式

在理解解码器运行的过程中，很容易联想一个标识下一个字符显示位置的不可见光标。但是请大家注意，如果您正在设计解码器，您应该获取最新的CEA-608规范，确保运行正确，因为本节只是一个简单的概述。

❑ 文本模式

基于实时滚动的文本模式使用显示器的7~15排进行显示，允许接收“恢复文本显示（Resume Text Display）”或者“文本重新启动（Text Restart）”代码。当选择文本模式时，文本存储器是空的，光标位于最顶端那一排的第1个字符位置上。一旦所有排都显示了文本，则启动文本滚动。

当接收到回车信号时，最顶端的那一排文本就被擦除，剩余文本平滑地向上滚动一行（通过12~26场使用6~13个均匀步），最底部那一排文本也被擦除，光标移到最底部那一排文本的第1个字符位置上。如果在滚动时接收到新的文本，就可以看到文本在显示区域从底部向上滚动。如果在滚动过程中接收到回车信号，那么各排立即向上移动一排，到达它们的最终位置上。

349 一旦光标移动到某排上的第32个字符上，在回车、前导地址码（preamble address code）或者退格之前接收到的任何文本，都将显示在第32个字符位置处，取代这个位置上以前的字符。“文本重新启动”命令将擦除显示器上的所有字符，并将光标移向首排第1个字符的位置上。

解码器还可以有选择地实现额外的实时显示方法，并且可以通过观众控制使用。

❑ 字幕模式

字幕模式有几种可用的模式，包括卷积（roll-up）、弹出（pop-on）和描绘（paint-on）模式。

卷积字幕通过接收一个杂项控制码（miscellaneous control code），从而选择字幕显示的排号来启用该模式。“卷积字幕，2排”启用第14排和第15排显示字幕；“卷积字幕，3排”启用第13~15排显示字幕；“卷积字幕，4排”启用第12~15排。无论启用了多少排来显示字幕，光标一直位于第15排上。一旦第15排满了，这一排文本就向上滚动一排（速度为1点/帧），而光标又移回到第15排的第1个字符上。

弹出字幕模式可以使用第1~4排或者第12~15排显示字幕，该模式通过“恢复字幕加载（Resume Caption Loading）”命令启用。显示存储器实质上是双缓冲的。当显示存储器缓冲器1时，存储器缓冲器2就加载字幕数据。在接收到“字幕结束（End of Caption）”代码时，就显示存储器缓冲器2，同时存储器缓冲器1加载新的字幕数据。

350 描绘字幕模式通过“恢复直接字幕（Resume Direct Captioning）”命令启用，与弹出字幕模式

相似，但是该模式不使用双缓冲策略；字幕数据直接加载到显示存储器中。

字符的格式、位置和属性使用3种类型的控制码[前导地址码、排中码（midrow code）和杂项控制码]指定。每个控制码包含两个字节，一起在第21行或第284行上传送。在第21行上，通常连续传送它们两次，以确保正确接收。但是在第284行上没有传送两次，从而使字幕使用的带宽最小。

控制码的第一个字节是非显示控制字节，范围从0x10到0x1F，第二个字节是一个显示控制字节，范围从0x20到0x7F。在每排的开头，发送一个控制码对这一排进行初始化。字幕卷积和文本模式允许在每排的开头使用前导地址码或者排中码，在其他字幕模式中使用前导地址码对这一排进行初始化。前导地址码如图8-55和表8-28所示。

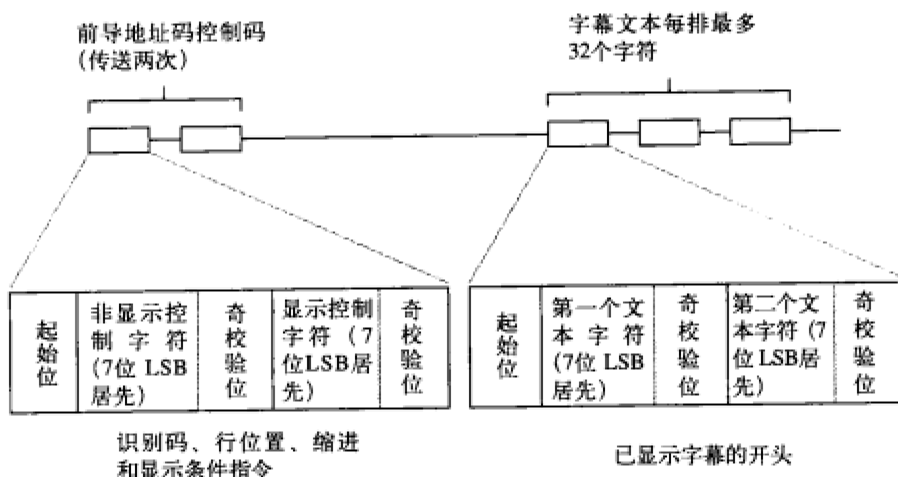


图8-55 隐藏字幕前导地址码格式

表8-28 隐藏字幕前导地址码。在文本模式中，缩进码可以用于执行缩进，在该实例中，忽略了排信息

非显示控制字节							显示控制字节							排位置
D6	D5	D4	D3	D2	D1	D0	D6	D5	D4	D3	D2	D1	D0	
0	0	1	CH	0	0	1	1	0	A	B	C	D	U	1
							1	1						2
				0	1	0	1	0						3
							1	1						4
				1	0	1	1	0						5
							1	1						6
				1	1	0	1	0						7
							1	1						8
				1	1	1	1	0						9
							1	1						10
				0	0	0	1	0						11
				0	1	1	1	0						12
							1	1						13
				1	0	0	1	0						14
							1	1						15

注：1. U：0＝没有下划线，1＝有下划线。
2. CH：0＝数据通道1，1＝数据通道2。

（续）

A	B	C	D	属　　性
0	0	0	0	白色
0	0	0	1	绿色
0	0	1	0	蓝色
0	0	1	1	青色
0	1	0	0	红色
0	1	0	1	黄色
0	1	1	0	洋红
0	1	1	1	白色斜体
1	0	0	0	缩进0，白色
1	0	0	1	缩进4，白色
1	0	1	0	缩进8，白色
1	0	1	1	缩进12，白色
1	1	0	0	缩进16，白色
1	1	0	1	缩进20，白色
1	1	1	0	缩进24，白色
1	1	1	1	缩进28，白色

排中码通常用在排的中间，用来改变字幕的颜色、斜体、下划线和闪烁属性，而且只能发生在字与字之间。颜色、斜体和下划线属性由前导地址码和排中码控制，闪烁属性由杂项控制码控制。直至接收到另一个控制码，或者到达这一排的末尾，否则属性一直有效。每一排以一个控制码开头，这个控制码用来设置颜色和下划线属性（如果在空白排的第一个字符之前没有接收到控制码，默认为白色没有下划线）。颜色属性只能由另一个颜色的排中码改变，斜体属性的改变不会改变颜色属性。但是，颜色属性可以关闭斜体属性。闪烁命令不会修改颜色、斜体或者下划线属性的状态。但是，颜色或者斜体排中控制码会关闭闪烁属性。注意下划线的颜色与被加上下划线的字符的颜色相同；下划线位于排的第11点上，包含字符列的整个宽度。

351

表8-29、图8-56和表8-30描述了排中和杂项控制码运算指令。例如，如果位于字幕末尾，那么控制码可能是“字幕结束”（传送两次）命令。它的后面可能是一个前导地址码（传送两次），启动另一行字幕。

表8-29　隐藏字幕排中码

<table>
<tr><th colspan="7">非显示控制字节</th><th colspan="7">显示控制字节</th><th rowspan="2">属　　性</th></tr>
<tr><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th></tr>
<tr><td rowspan="8">0</td><td rowspan="8">0</td><td rowspan="8">1</td><td rowspan="8">CH</td><td rowspan="8">0</td><td rowspan="8">0</td><td rowspan="8">1</td><td rowspan="8">0</td><td rowspan="8">1</td><td rowspan="8">0</td><td>0</td><td>0</td><td>0</td><td rowspan="8">U</td><td>白色</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>绿色</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>蓝色</td></tr>
<tr><td>0</td><td>1</td><td>1</td><td>青色</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>红色</td></tr>
<tr><td>1</td><td>0</td><td>1</td><td>黄色</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>洋红</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>斜体</td></tr>
</table>

注：1. U：0＝没有下划线，1＝有下划线。
2. CH：0＝数据通道1，1＝数据通道2。
3. 斜体实现为字符的垂直范围向右倾斜两点。一些解码器实现为每4行扫描行倾斜一个点。下划线位于排的第22点和第23点上，包含整个列宽。

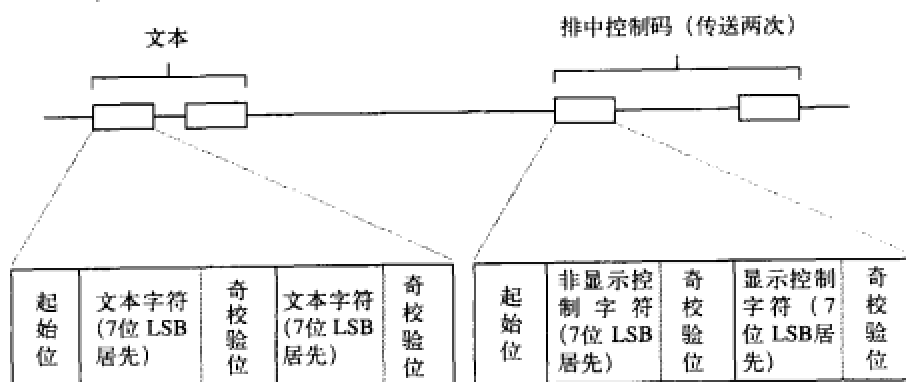


图8-56 隐藏字幕排中码格式。杂项控制码也可以在排中控制码的位置中传送

表8-30 隐藏字幕杂项控制码

<table>
<tr><th colspan="7">非显示控制字节</th><th colspan="7">显示控制字节</th><th rowspan="2">命 令</th></tr>
<tr><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th></tr>
<tr><td rowspan="16">0</td><td rowspan="16">0</td><td rowspan="16">1</td><td rowspan="16">CH</td><td rowspan="16">1</td><td rowspan="16">0</td><td rowspan="16">F</td><td rowspan="16">0</td><td rowspan="16">1</td><td rowspan="16">0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>恢复字幕加载</td></tr>
<tr><td>0</td><td>0</td><td>0</td><td>1</td><td>退格</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>0</td><td>保留</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>1</td><td>保留</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>0</td><td>删除到排尾</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>1</td><td>卷积字幕，2排</td></tr>
<tr><td>0</td><td>1</td><td>1</td><td>0</td><td>卷积字幕，3排</td></tr>
<tr><td>0</td><td>1</td><td>1</td><td>1</td><td>卷积字幕，4排</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>0</td><td>闪烁</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>1</td><td>恢复直接字幕</td></tr>
<tr><td>1</td><td>0</td><td>1</td><td>0</td><td>文本重新启动</td></tr>
<tr><td>1</td><td>0</td><td>1</td><td>1</td><td>恢复文本显示</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>0</td><td>擦除已显示存储器</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>1</td><td>回车</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>擦除未显示存储器</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>1</td><td>字幕结束（交换存储器）</td></tr>
<tr><td rowspan="3">0</td><td rowspan="3">0</td><td rowspan="3">1</td><td rowspan="3">CH</td><td rowspan="3">1</td><td rowspan="3">1</td><td rowspan="3">1</td><td rowspan="3">0</td><td rowspan="3">1</td><td rowspan="3">0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>制表偏移（1列）</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>0</td><td>制表偏移（2列）</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>1</td><td>制表偏移（3列）</td></tr>
</table>

注：1. F：0＝第21行，1＝第284行。CH：0＝数据通道1，1＝数据通道2。
2. 与“闪烁”属性关联的字符每秒闪烁一次，每次0.25秒。

字符使用点矩阵格式显示。每个字符单元通常占用16样本宽和26样本高（16×26），如图8-57所示。每排的第2~19点通常用于实际字符轮廓。每排的第0、1、20、21、24和25点通常是空白的，以提供字符之间的垂直空间，而且下划线通常位于每排的第22点和第23点上。每列的第0、1、14和15点是空白的，以提供字符之间的水平空间，但是当显示下划线时，每排的第22点和第23点除外。这样就可以在字符ROM中存储12×18个字符。表8-31示范了基本的字符集。



图8-57　第1排的典型16×26隐藏字幕字符单元格式

表8-31　隐藏字幕基本字符集

非显示控制字节							显示控制字节							特殊字符
D6	D5	D4	D3	D2	D1	D0	D6	D5	D4	D3	D2	D1	D0	
0	0	1	CH	0	0	1	0	1	1	0	0	0	0	®
										0	0	0	1	°
										0	0	1	0	1/2
										0	0	1	1	¿
										0	1	0	0	TM
										0	1	0	1	¢
										0	1	1	0	£
										0	1	1	1	音乐注释
										1	0	0	0	à
										1	0	0	1	透明空格
										1	0	1	0	è
										1	0	1	1	â
										1	1	0	0	ê
										1	1	0	1	î
										1	1	1	0	ô
										1	1	1	1	û

（续）

D6D5D4D3	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
D2D1D0												
000		(	0	8	@	H	P	X	ú	h	p	x
001	!	)	1	9	A	I	Q	Y	a	i	q	y
010	“	á	2	:	B	J	R	Z	b	j	r	z
011	#	+	3	;	C	K	S	[	c	k	s	ç
100	$	,	4	<	D	L	T	é	d	l	t	÷
101	%	-	5	=	E	M	U	]	e	m	u	Ñ
110	&	.	6	>	F	N	V	í	f	n	v	ñ
111	'	/	7	?	G	O	W	ó	g	o	w	

一些字幕解码器支持多种位于16×26区域内的字符大小，包括13×16、13×24、12×20和12×26。但是，由于可用的有限显示区域，并不是所有的组合都能生成一个切合实际的结果。

3. 可选字幕特征

高级字幕解码器还有3组可选的特征。

❑ 可选属性

高级字幕解码器还提供有额外的颜色选项，如表8-32所示。

表8-32 隐藏字幕可选属性代码

非显示控制字节							显示控制字节							背景属性
D6	D5	D4	D3	D2	D1	D0	D6	D5	D4	D3	D2	D1	D0	
0	0	1	CH	0	0	0	0	1	0	0	0	0	T	白色
										0	0	1		绿色
										0	1	0		蓝色
										0	1	1		青色
										1	0	0		红色
										1	0	1		黄色
										1	1	0		洋红
										1	1	1		黑色
0	0	1	CH	1	1	1	0	1	0	1	1	0	1	透明
D6	D5	D4	D3	D2	D1	D0	D6	D5	D4	D3	D2	D1	D0	前景属性
0	0	1	CH	1	1	1	0	1	0	1	1	1	0	黑色
													1	黑下划线

注：1. F：0＝不透明，1＝半透明。

2. CH：0＝数据通道1，1＝数据通道2。

3. 下划线位于每排的第22点和第23点上，包含整个列宽。

352
~
353

如果解码器不支持半透明颜色，可以使用不透明颜色。如果解码器不支持某个特定的背景颜色，默认为黑色背景颜色。但是，如果解码器支持黑色背景颜色，那么解码器应该实现所有的背景颜色。

背景属性看上去就像显示器上的标准空格，而且直至排结束或者接收到另一个背景属性，否则背景属性保持不变。

前景属性提供一种第八种颜色（黑色）作为字符颜色。和排中码一样，前景属性代码关闭斜

体和闪烁属性，而且最低有效位控制下划线。

背景和前景属性代码拥有一个自动的退格，从而支持当前解码器向后兼容。因此，属性前面必须加上一个标准的空格字符。标准解码器显示空格并忽略该属性。扩展的解码器显示空格，并且在接收到属性时，退后一格，然后显示一个改变了颜色和不透明度的空格。因此，无论解码器是什么类型，文本格式都是相同的。

❑ 可选隐藏组扩展

354～357

为了支持当前标准还没有定义的新特征和字符，CEA维护了一组由不同字幕提供商和解码器制造商请求的代码分配。这些代码分配（目前用于选择不同的亚洲字符集）与美国的字幕解码器不兼容，使用它们的视频不应该在美国市场上发布。

隐藏组扩展需要2个字节。表8-33列举了支持亚洲语言字幕的隐藏组扩展的当前分配情况。

表8-33　隐藏字幕可选隐藏组扩展

<table>
<tr><th colspan="7">非显示控制字节</th><th colspan="7">显示控制字节</th><th rowspan="2">背景属性</th></tr>
<tr><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th><th>D6</th><th>D5</th><th>D4</th><th>D3</th><th>D2</th><th>D1</th><th>D0</th></tr>
<tr><td rowspan="7">0</td><td rowspan="7">0</td><td rowspan="7">1</td><td rowspan="7">CH</td><td rowspan="7">1</td><td rowspan="7">1</td><td rowspan="7">1</td><td rowspan="7">0</td><td rowspan="7">1</td><td rowspan="7">0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>标准字符集（普通大小）</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>1</td><td>标准字符集（双字节大小）</td></tr>
<tr><td>0</td><td>1</td><td>1</td><td>0</td><td>第一个专有字符集</td></tr>
<tr><td>0</td><td>1</td><td>1</td><td>1</td><td>第二个专有字符集</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>0</td><td>中华人民共和国字符集（GB2312）</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>1</td><td>韩国标准字符集（KSC 5601-1987）</td></tr>
<tr><td>1</td><td>0</td><td>1</td><td>0</td><td>第一个注册字符集</td></tr>
</table>

注：CH：0＝数据通道1；1＝数据通道2。

❑ 可选扩展字符

解码器还可以支持另外64个重音字符（8个字符集，每个字符集有8个字符），允许显示其他语言（例如西班牙语、法语、葡萄牙语、德语、丹麦语、意大利语、芬兰语和瑞典语）。如果支持，这些重音字符就可以在所有的字幕和文本模式中使用。

每个扩展的字符都集成一个自动的退格，从而支持当前解码器向后兼容。因此，扩展的字符必须在前面加上标准的ASCII字符。标准解码器显示ASCII字符，并忽略重音字符。扩展的解码器显示ASCII字符，在接收到重音字符时，退后一格，然后显示重音字符。因此，无论解码器是什么类型，文本格式都是相同的。

扩展的字符需要2个字节。第一个字节是数据通道1，值为0x12或者0x13（数据通道2的值为0x1A或0x1B），后面跟一个位于0x20~0x3F之间的值。

4. 扩展数据服务

在带宽可用时，第284行可以包含扩展数据服务信息，与字幕和文本信息隔行扫描。在这种情况下，控制码不用再传送2次，因为它们可能是针对字幕和文本服务的。 358

信息作为包传送，并作为一个独立的独特数据通道。每个包的数据可以是连续的，也可以是不连续的，而且还可以进一步分成子包（subpacket），从而可以插入到第284行信息流中的任何可用空间中。

扩展数据字符有以下4种类型。

控制：控制字符用作一种模式开关，启用扩展数据模式。控制字符是两个字符中的第一个字符，值的范围为0x01~0x0F。

类型：类型字符在控制字符之后（因此，它们是两个字符中的第二个字符），用来指定包的类型。类型字符值的范围为0x01~0x0F。

校验和：校验和字符总是位于“包结束”控制字符之后。因此，校验和字符是两个字符中的第二个字符，值的范围从0x00到0x7F。

信息：这些字符可能是ASCII或非ASCII数据。它们成对传送，包含32个字符。其中NUL字符（0x00）用来确保字符总是成对发送。

❑ 控制字符

表8-34列举了各种控制码。“当前”类别描述当前正在传送的节目。“未来节目”类别描述稍后传送的节目。它包含的信息和格式与“当前”类别相同。“通道信息”类别描述与节目不相关的有关通道的信息。“杂项”类别描述杂项信息。“公共服务”类别传送公共服务类型的数据或者消息。“私有数据”类别用于电视制式希望拥有的专有系统。

表8-34 CEA-608控制码

控制码	功能	类别
0x01	启动	
0x02	继续	当前
0x03	启动	
0x04	继续	未来
0x05	启动	通道信息
0x06	继续	
0x07	启动	杂项
0x08	继续	
0x09	启动	公共服务
0x0A	继续	
0x0B	启动	保留
0x0C	继续	
0x0D	启动	私有数据
0x0E	继续	
0x0F	结束	所有

❑ 类型定义（当前类别和未来节目）

■ 节目标识号（0x01）

这个包使用4个字符来指定一个相对于格林威治标准时间（Coordinated Universal Time，UTC）的预定启动时间和日期。节目标识号的格式如表8-35所示。

表8-35 CEA-608节目标识号格式

D6	D5	D4	D3	D2	D1	D0	字 符
1	m5	m4	m3	m2	m1	m0	分钟
1	D	h4	h3	h2	h1	h0	小时
1	L	d4	d3	d2	d1	d0	日
1	Z	T	m3	m2	m1	m0	月

分钟的范围为0~59。小时的范围为0~23。日的范围为1~31。月的范围为1~12。T表示节目针对山地和太平洋时区（Mountain and Pacific time zone）是否例行磁带延迟。解码器忽略D、L和Z位。当所有的字符都是1时，它表示当前节目的尾部。

■长度/播放时间（0x02）

这个包有2、4或6个字符，表示节目的预定长度和节目的已播放时间。长度/播放时间的格式如表8-36所示。

表8-36 CEA-608长度/播放时间格式

D6	D5	D4	D3	D2	D1	D0	字 符
1	m5	m4	m3	m2	m1	m0	长度，分钟
1	h5	h4	h3	h2	h1	h0	长度，小时
1	m5	m4	m3	m2	m1	m0	已播放时间，分钟
1	h5	h4	h3	h2	h1	h0	已播放时间，小时
1	s5	s4	s3	s2	s1	s0	已播放时间，秒
0	0	0	0	0	0	0	空字符

分钟和秒的范围为0~59。小时的范围为0~63。

■节目名称（0x03）

这个包包含2~32个ASCII字符，指定节目的标题。

■节目类型（0x04）

这个包包含2~32个字符，指定节目的类型。每个字符都编码为一个关键字，如表8-37所示。

表8-37 CEA-608节目类型

代码（十六进制）	关 键 字	代码（十六进制）	关 键 字	代码（十六进制）	关 键 字
20	教育	30	商业	40	幻想
21	娱乐	31	经典	41	农业
22	电影	32	大学	42	时尚
23	新闻	33	战斗	43	虚构
24	宗教	34	喜剧	44	食品
25	体育	35	解说	45	足球
26	其他	36	音乐会	46	外语
27	动作	37	消费	47	投资
28	广告	38	现代	48	游戏/测验
29	动画	39	犯罪	49	园艺
2A	诗歌	3A	舞蹈	4A	高尔夫
2B	汽车	3B	纪录片	4B	政府
2C	颁奖	3C	戏剧	4C	健康
2D	棒球	3D	小学	4D	高中

（续）

代码（十六进制）	关　键　字	代码（十六进制）	关　键　字	代码（十六进制）	关　键　字
2E	篮球	3E	成人	4E	历史
2F	通报	3F	锻炼	4F	业余爱好
50	曲棍球	60	音乐	70	传奇
51	家庭	61	探秘	71	科学
52	恐怖	62	民族	72	系列片
53	信息	63	自然	73	服务
54	指南	64	警察	74	购物
55	国际	65	政治	75	肥皂剧
56	会谈	66	公演	76	特别节目
57	语言	67	预先录制	77	悬念
58	法律	68	产品	78	谈话
59	生活	69	职业	79	技术
5A	本地新闻	6A	公共	7A	乒乓球
5B	数学	6B	竞赛	7B	旅游
5C	医疗	6C	阅读	7C	杂技
5D	会议	6D	修理	7D	录像
5E	军事	6E	重播	7E	天气
5F	电视连续剧	6F	评论	7F	西方

■内容分级（content advisory）（0x05）

这个包，就是通常所指的“V-chip[①]”，包含表8-38中所示的信息，表示节目的分级。

表8-38　CEA-608内容分级格式

D6	D5	D4	D3	D2	D1	D0
1	D/a2	a1	a0	r2	r1	r0
1	V/FV	S	L/a3	g2	g1	g0

r2~r0：MPA分级
- 000　不可用
- 001　G
- 010　PG
- 011　PG-13
- 100　R
- 101　NC-17
- 110　X
- 111　没有分级

g2~g0：美国电视分级
- 000　没有等级
- 001　TV-Y
- 010　TV-Y7
- 011　TV-G
- 100　TV-PG
- 101　TV-14
- 110　TV-MA
- 111　没有分级

a3~a0：xxx0 MPA分级
- LD01　美国电视父母引导
- 0011　加拿大英语语言分级
- 0111　加拿大法语语言分级
- 1011　保留
- 1111　保留

g2~g0：加拿大英语语言分级
- 000　E
- 001　C
- 010　C8+
- 011　G
- 100　PG
- 101　14+
- 110　18+
- 111　保留

g2~g0：加拿大法语语言分级
- 000　E
- 001　G
- 010　8ans+
- 011　13ans+
- 100　16ans+
- 101　18ans+
- 110　保留
- 111　保留

FV表示节目内容存在幻想式的暴力。V表示节目内容存在暴力形为。S表示节目中有性相关

① 一种暴力节目过滤芯片。——译者注

的状况。L表示节目内容存在强烈粗俗语言。D表示节目内容存在性暗示的对话。

■音频服务（0x06）

这个包包含2个表8-39中所示的字符，表示可用的音频语言和类型。

表8-39 CEA-608音频服务格式

D6	D5	D4	D3	D2	D1	D0	字 符
1	L2	L1	L0	T2	T1	T0	主音频节目
1	L2	L1	L0	S2	S1	S0	辅助音频节目（SAP）

L2~L0：	T2~T0：	S2~S0：
000 未知	000 未知	000 未知
001 英语	001 单声道	001 单声道
010 西班牙语	010 模拟立体声	010 视频描述
011 法语	011 真立体声	011 非节目音频
100 德语	100 立体声环绕	100 特殊效果
101 意大利语	101 数据服务	101 数据服务
110 其他	110 其他	110 其他
111 无	111 无	111 无

■字幕服务（0x07）

这个包包含2~8个表8-40中所示的字符，表示可用的节目字幕服务。L2~L0的编码如表8-39所示。

表8-40 CEA-608字幕服务格式

D6	D5	D4	D3	D2	D1	D0	字 符
1	L2	L1	L0	F	C	T	服务代码

FCT：000 第21行，数据通道1字幕
001 第21行，数据通道1文本
010 第21行，数据通道2字幕
011 第21行，数据通道2文本
100 第284行，数据通道1字幕
101 第284行，数据通道1文本
110 第284行，数据通道2字幕
111 第284行，数据通道2文本

■复制和再发行控制包（0x08）

这个CGMS-A（Copy Generation Management System-Analog，复制代次管理系统-模拟）和再发行控制描述符（Redistribution Control Descriptor，RCD）包包含2个表8-41中所示的字符。

表8-41 CEA-608复制和再发行控制包格式

D6	D5	D4	D3	D2	D1	D0
1	0	B4	B3	B2	B1	B0
0	0	0	0	0	0	RCD

B4~B3 CGMS-A服务：	B2~B1 模拟保护服务（APS）：
00 允许无限制的复制	00 无模拟保护服务
01 不允许再复制	01 打开PSP（pseudo-sync pulse，伪同步脉冲），关闭彩色分条
10 允许一代复制	10 打开PSP，打开2行彩色分条
11 不允许复制	11 打开PSP，打开4行彩色分条

无论B3还是B4为0，都没有模拟保护服务（Analog Protection Service，APS）（B1和B2都为0）。B0是模拟源位。

当RCD是1时，就会以某种形式表示消费者再发行的控制，例如提供ATSC再发行控制描述符。

■复合包-1（0x0C）

这个包是一种将若干个包作为一个组传送的方式。该包包含节目类型（5个字符）、内容分级（1个字符）、长度（2个字符）、播放时间（2个字符）和节目名称（0~22个字符）。

■复合包-2（0x0D）

这个包是一种将若干个包作为一个组传送的方式。该包包含节目ID（4个字符）、音频服务（2个字符）、字幕服务（2个字符）、电码字符（Call Letter）（4个字符）、本地频道（Native Channel）（2个字符）和网络名称（0~18个字符）。

■节目描述排1至节目描述排8（0x10~0x17）

这个包包含1~8个包排，每个包排包含0~32个ASCII字符。一个包排对应显示器上的一行文本。每个包以数字序列使用，如果包中不包含ASCII字符，将显示一个空白行。

❑类型定义（通道信息类）

■网络名称（0x01）

这个包使用2~32个ASCII字符来指定网络名称。

■电码字符和本地频道（0x02）

这个包使用4个或者6个ASCII字符来指定频道的电码字符。当使用6个字符时，它们反映FCC分配的无线频道号（2~69）。单个数字的频道号前面加一个0或者一个空字符。

■磁带延迟（0x03）

这个包使用2个字符来指定地方台通常延迟网络节目的小时和分钟数量。这个包的格式如表8-42所示。

表8-42 CEA-608磁带延迟格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	m5	m4	m3	m2	m1	m0	分钟
1	-	h4	h3	h2	h1	h0	小时

分钟的范围为0~59。小时的范围为0~23。这个延迟应用到频道上所有在节目ID包（表8-35）中设置了T位的节目。

■传输信号标识符（0x04）

这个包包含4个字符，传送唯一的16位传输信号标识符（Transmission Signal Identifier，TSID），传输信号标识符分配给发端模拟许可证。这个包的格式如表8-43所示。

360~362

表8-43 CEA-608传输信号标识符（TSID）格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	–	–	t3	t2	t1	t0	TSID 0
1	–	–	t7	t5	t5	t4	TSID 1
1	–	–	t11	t10	t9	t8	TSID 2
1	–	–	t15	t14	t13	t12	TSID 4

■时间（0x01）

这个包使用6个字符指定相对于格林威治时间（UTC）的当前星期、月和日。这个包的格式

如表8-44所示。

表8-44　CEA-608时间格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	m5	m4	m3	m2	m1	m0	分钟
1	D	h4	h3	h2	h1	h0	小时
1	L	d4	d3	d2	d1	d0	日
1	Z	T	m3	m2	m1	m0	月
1	-	-	-	D2	D1	D0	星期
1	Y5	Y4	Y3	Y2	Y1	Y0	年

分钟的范围为0~59。小时的范围为0~23。日的范围为1~31。月的范围为1~12。星期的范围为1（星期日）到7（星期六）。年的范围为0~63（加上1990）。

T表示节目针对山地和太平洋时区是否例行磁带延迟。D表示当前是否遵守夏令时时间。L表示当UTC时间是3月1日时，本地时间是2月28日还是2月29日。Z表示秒是否设置为0（从而允许不必传送完整的6bit的秒数据来协调时间）。

■脉冲捕捉ID（0x02）

这个包携带节目的起始时间和长度，可以用来告诉VCR录制该节目。这个包的格式如表8-45所示。

表8-45　CEA-608脉冲捕捉ID格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	m5	m4	m3	m2	m1	m0	启动时间，分钟
1	D	h4	h3	h2	h1	h0	启动时间，小时
1	L	d4	d3	d2	d1	d0	启动时间，日
1	Z	T	m3	m2	m1	m0	启动时间，月
1	m5	m4	m3	m2	m1	m0	长度，分钟
1	h5	h4	h3	h2	h1	h0	长度，小时

起始和长度分钟的范围为0~59。起始小时的范围为0~23，长度小时的范围为0~63。日的范围为1~31。月的范围为1~12。T表示节目针对山地和太平洋时区是否例行磁带延迟。解码器忽略D、L和Z位。

■补充数据位置（0x03）

这个包使用2~32个字符来指定额外VBI数据所在的其他行。表8-46示范了这个包的格式。

表8-46　CEA-608补充数据位置格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	F	N4	N3	N2	N1	N0	位置

F指示场1（0）或者场2（1）。N的值的范围为7~31，表示一个具体的行号。

■本地时区和DST应用（0x04）

这个包使用2个字符指定观众的时区及当地是否采用夏令时。这个包的格式如表8-47所示。

表8-47　CEA-608本地时区和DST应用格式

D6	D5	D4	D3	D2	D1	D0	字　符
1	D	h4	h3	h2	h1	h0	小时
0	0	0	0	0	0	0	空

小时的范围为0~23。这是相对于UTC的标称时区偏移，单位是小时。当这个地区使用夏令时时，D为1。

■ 带外频道号（0x40）

这个包使用2个字符指定一个所有随后的带外包访问的频道号。这是后面的带外包所属的CATV频道号。这个包的格式如表8-48所示。

表8-48 CEA-608带外频道号格式

D6	D5	D4	D3	D2	D1	D0	字 符
1	c5	c4	c3	c2	c1	c0	低频道
1	c11	c10	c9	c8	c7	c6	高频道

■ 频道映射指针（0x41）这个包使用2个字符指定包含频道映射头（Channel Map Header）和频道映射包的频道号。

■ 频道映射头包（0x42）这个包使用4个字符指定频道映射的频道号和当前映射的版本号。

■ 频道映射包（0x43）这个包使用2个或者4个字符指定用户频道号以及相应的调谐器频道号。最多可以包含6个可选的隐藏字幕字符来传送用户频道的电码字符或者网络ID。

363~366

❑ 类型定义（公共服务类）

■ 国家天气服务代码（0x01）

■ 这个包传送与天气相关的紧急广播消息，消息指示状况的类别、受影响国家以及截止时间。

■ 国家天气服务消息（0x02）

这个包可以传送多达32个字符的实际文本消息，该文本消息由国家天气服务发送。

5. 字幕（CC）和文本（T）通道

CC1、CC2、T1和T2都位于第21行上。CC3、CC4、T3和T4都位于第284行上。第284行上的第5通道携带扩展数据服务。T1~T4与CC1~CC4相似，但是在显示滚动文本信息时要占用整个或者一半屏幕。

CC1通常用于主字幕通道。CC2或者CC3偶尔用来支持第二语言。

6. PAL制式隐藏字幕

对于（M）PAL制式，字幕数据可以在第18行和第281行上提供，但是，由于视频编辑，字幕数据偶尔可能会位于第18行至第22行之间的任意一行以及第281行至第285行之间的任意一行上。

对于（B、D、G、H、I、N、NC）PAL制式录像带，字幕数据可能位于第22行和第335行上，但是，由于编辑，字幕数据偶尔可能会位于第22行至第26行之间的任意一行以及第335行至第339行之间的任意一行上。PAL制式隐藏字幕的数据格式、幅度和上升及下降时间与美国所使用的隐藏字幕匹配。但是，由于PAL制式625行的水平时序，PAL制式隐藏字幕的时序与美国所使用的隐藏字幕稍微有所不同，PAL制式625行的时序如图8-58所示。

367~368

8.5.3 宽屏信令和CGMS

为了方便电视机接收到的不同宽高比节目素材的处理，开发了一种称之为宽屏信令（Widescreen Signaling，WSS）的系统。这一标准允许支持WSS的16：9电视机按节目原始宽高比显示节目。

1. 625i系统

625i（576i）系统的宽屏信令系统基于ITU-R BT.1119和ETSI EN 300 294标准。对于YPbPr和S-video接口，WSS在Y信号上提供。对于模拟RGB接口，WSS在全部3个信号上提供。

WSS信号也支持模拟复制代次管理系统（Analog Copy Generation Management System，CGMS-A）。

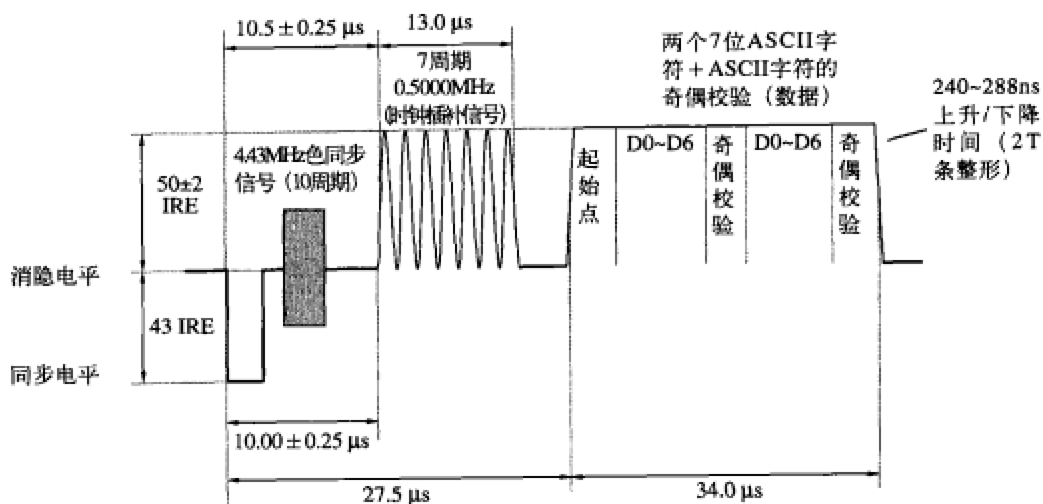


图8-58 625行系统第22行和第335行隐藏字幕时序

❑ 数据时序

对于（B、D、G、H、I、N、NC）PAL制式，WSS数据通常位于第23行上，如图8-59所示。但是，由于视频编辑，WSS数据可能位于第23行至第27行之间的任意一行上。

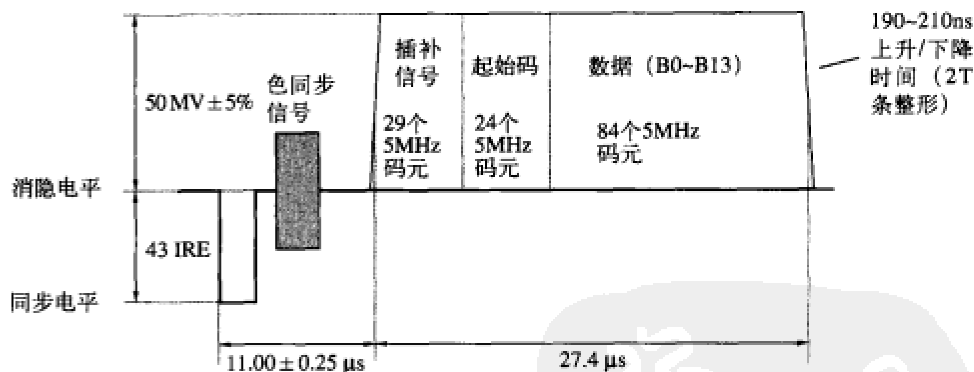


图8-59 625行系统第23行WSS时序

WSS数据的时钟频率为5MHz（±100Hz）。它的信号波形是一个正弦形状的脉冲，半幅度宽度为200±10ns。信号幅度为500mV±5%。

NRZ数据位由双相代码调制器处理，这样一个数据周期等于6个5MHz码元。

❑ 数据内容

WSS包含一个插补信号码、一个起始码和14bit数据，如表8-49所示。

表8-49 625行系统WSS信息

插补信号	29个5MHz码元	1 1111 0001 1100 0111 0001 1100 0111 （0x1F1C 71C7）
起始码	24个5MHz码元	0001 1110 0011 1100 0001 1111 （0x1E 3C1F）
组A（宽高比）	24个5MHz码元 0表示000 111 1表示111 000	b0，b1，b2，b3

（续）

组B（增强服务）	24个5MHz码元 0表示000 111 1表示111 000	b4，b5，b6，b7 （因为b7保留，所以b7＝0）
组C（字幕）	18个5MHz码元 0表示000 111 1表示111 000	b8，b9，b10
组D（保留）	18个5MHz码元 0表示000 111 1表示111 000	b11，b12，b13

插补信号包含29个特定序列的5MHz码元，如表8-49所示。

起始码包含24个特定序列的5MHz码元，如表8-49所示。

组A数据包含4个数据位，用来指定宽高比。每个数据位生成6个5MHz码元。b0是LSB。 369

表8-50列举了组A数据所包含数据位的分配和使用情况。表8-50中列举的有效行的号码是准确的宽高比（a＝1.33、1.56或1.78）。

表8-50 625行系统WSS信息组A（宽高比）数据位分配和使用情况

b0，b1，b2，b3	宽高比标签	格 式	4：3显示中的位置	有效行数量	最低要求
0001	4：3	完全模式	–	576	情况1
1000	14：9	Letterbox模式	居中	504	情况2
0100	14：9	Letterbox模式	置顶	504	情况2
1101	16：9	Letterbox模式	居中	430	情况3
0010	16：9	Letterbox模式	置顶	430	情况3
1011	>16：9	Letterbox模式	居中	–	情况4
0111	14：9	完全模式	居中	576	–
1110	16：9	完全模式（变形）	–	576	–

宽高比标签指示一组可能的宽高比（a）范围以及有效行的数量：

4：3	a≤1.46	527~576
14：9	1.46<a≤1.66	463~526
16：9	1.66<a≤1.90	405~462
>16：9	a>1.90	<405

为了允许自动选择显示器模式，16：9接收机应该支持下述最低要求。

情况1：4：3宽高比的图像应该在显示器上居中显示，左右边没有图像的部分显示黑条。

情况2：14：9宽高比的图像应该在显示器上居中显示，左右边没有图像的部分显示黑条。还有一种方案，即图像通过一个较小的（通常为8%）水平几何失真，以显示器完整宽度进行显示。

情况3：16：9宽高比的图像应该以显示器的完整宽度进行显示。

情况4：>16：9宽高比的图像应该以情况3进行显示，或者通过缩小以显示器的完整高度进行显示。

■ 组B数据

组B数据包含4个数据位，用来指定增强服务。每个数据位生成6个5MHz码元。数据位b4是LSB。b5和b6位用于PALplus。

b4：模式。0：摄像机模式1：影片模式。

b5：彩色编码。0：普通PAL 1：运动自适应ColorPlus。

b6：辅助信号。0：没有辅助信号1：有辅助信号。

■ 组C数据

组C数据包含3个数据位，用来指定字幕。每个数据位生成6个5MHz码元。数据位b8是LSB。

b8：图文电视字幕。0：没有。1：有。

b9，b10：开放字幕。00：没有。01：有效图像外。10：有效图像内。11：保留。

■ 组D数据

组D数据包含3个数据位，用来指定环绕音频和复制保护。每个数据位生成6个5MHz码元。数据位b11是LSB。

370~371

b11：环绕音频。0：没有。1：有。

b12：版权。0：未声明版权或者未知。1：声明版权。

b13：复制保护。0：不限制复制。1：限制复制。

2. 525i系统

EIA-J CPR-1204和IEC 61880为525i（480i）系统定义了宽屏信令标准。对于YPbPr和S-video接口，WSS在Y信号上提供。对于模拟RGB接口，WSS在全部3个信号上提供。

❑ 数据时序

第20行和第283行用来传送WSS信息，如图8-60所示。但是，由于视频编辑，WSS信息可能位于第20至第24行之间的任意一行以及第283行至第287行之间的任意一行上。

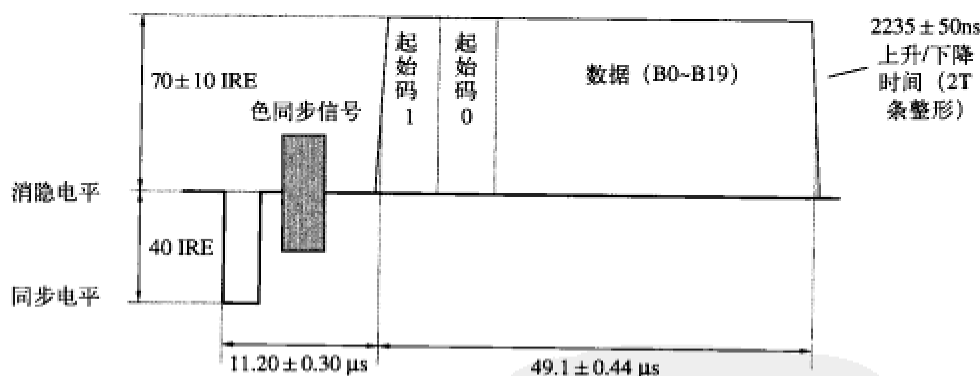


图8-60 525行系统第20行和第283行WSS时序

WSS信息的时钟频率为$F_{SC}/8$或者大约447.443kHz。F_{SC}是3.579 545MHz的彩色副载波频率。它的信号波形是一个正弦形状的脉冲，半幅度宽度为2.235μs±50ns。1的信号幅度为70±10 IRE，0的信号幅度为0±5 IRE。

❑ 数据内容

WSS包含2bit起始码、14bit数据和6bitCRC，如表8-51所示。所使用的CRC为X^6+X+1，所有位都预置为1。

表8-51 525行系统WSS信息数据位分配和使用情况

起始码	1
起始码	0
字0	b0, b1
字1	b2, b3, b4, b5
字2	b6, b7, b8, b9, b10, b11, b12, b13
CRC	b14, b15, b16, b17, b18, b19

■ 起始码

起始码包含一个1数据位，后面跟一个0数据位，如表8-51所示。

■ 字0数据

字0数据包含2个数据位：

b1，b0：

00	4：3宽高比	普通
01	4：3宽高比	Letterbox模式
10	16：9宽高比	变形
11	保留	

■ 字1数据

字1数据包含4个数据位：

b2，b3，b4，b5：

0000	复制控制信息
1111	默认

当字1数据是0000时，复制控制信息在字2数据中传送。当没有传送复制控制信息时，字1数据必须设置为默认值1111。

■ 字2数据

字2数据包含14个数据位。当字1数据是0000时，字2数据包含复制控制信息。字2复制控制信息的传送速度必须大于等于每2秒2帧。

位b6和b7用一个模拟信号（CGMS-A）指定复制代次管理系统。CGMS-A包含2比特数字信息：

b6，b7：

00	允许复制
01	保留
10	只允许复制一次
11	不允许复制

这个CGMS-A信息通常也必须通过第284行的扩展数据服务复制和再发行控制包（在介绍隐藏字幕时讨论）传送。

位b8和b9指定添加在模拟NTSC或者PAL视频信号上的模拟保护服务（Analog Protection Service，APS）。

b8，b9：

00	没有模拟保护服务
01	打开PSP，关闭彩色分条
10	打开PSP，打开2行彩色分条
11	打开PSP，打开4行彩色分条

372

PSP是一个伪同步脉冲运算，如果打开，将提供在复合、S-video和（YPbPr的）Y模拟视频输出上。彩色分条运算将复合和S-video模拟视频输出上特定扫描行上的前半部分色同步信号的正常相位反转。

该模拟保护服务（Analog Protection Service，APS）通常也必须通过第284行的扩展数据服务复制和再发行控制包（在介绍隐藏字幕时讨论）传送。

位b10指定源是否来源于模拟预录制媒介（pre-recorded medium）。

b10：

0	不是模拟预录制媒介
1	模拟预录制媒介

位b11、b12和b13保留，值为000。

8.5.4 图文电视

图文电视可以传送文本、图形和数据。尽管图文电视主要通过使用VBI期传送，但是图文电视数据可以在任意行上传送。图文电视标准由ETSI ETS 300 706、ITU-R BT.653和EIA-516规定。

对于YPbPr和S-video接口，图文电视在Y信号上提供。对于模拟RGB接口，图文电视在全部3个信号上提供。

目前有许多系统利用图文电视物理层来传送专有信息。利用图文电视物理层的优点是图文电视已经在许多国家获批进行广播，因此不需要再认证一个新的传输技术。

图文电视的数据传送速率比隐藏字幕快许多，在一些情况下可以高达7Mbit/s。因此，需要使用重影消除可靠地恢复所传送的数据。

目前定义的图文电视系统有7种，如表8-52所示。系统B［也称为世界图文电视广播系统（World System Teletext，WST）］已成为图文电视实际上的标准，被广泛接受。

表8-52 图文电视系统及参数概括

参　数	系统A	系统B	系统C	系统D
625行系统视频系统				
速率（Mbit/s）	6.203125	6.9375	5.734375	5.6427875
数据幅度	67 IRE	66 IRE	70 IRE	70 IRE
每行数据	40B	45B	36B	37B
525行系统视频系统				
速率（Mbit/s）	–	5.727272	5.727272	5.727272
数据幅度	–	70 IRE	70 IRE	70 IRE
每行数据	–	37B	36B	37B

373~374

EIA-516，也称为NABTS（North American Broadcast Teletext Specification，北美广播图文电视规范）曾在美国有少量应用，它是BT.653 525行系统C标准的扩展。

图8-61描述了一个扫描行上的图文电视数据。如果某行包含有色同步信号，那么在图文电视数据存在的情况下色同步信号仍然存在。16bit时钟插补（或者时钟同步）信号由交互的1和0构成。

图8-62和图8-63分别描述了图文电视系统B和系统C的结构。

系统B图文电视概述

因为图文电视系统B是图文电视的事实标准，所以这里对它进行一个基本的概述。

图文电视服务通常由页（page）构成，每个页对应于一屏幕信息。页一次传送一页，当传送完所有页之后，又重新开始一个周期，通常一个周期的时间大约为30秒。但是，如果愿意，广播商可以多次传送某些页。

图文电视服务通常基于杂志（magazine），最多有8本杂志（即最多允许8个独立的图文电视服务），每本杂志所含页最多可达100页。杂志1所使用页的编号为100~199，杂志2所使用页的编号为200~299，等等，依此类推。每个页还可以有子页（sub-page），用来在一本杂志内扩展页的数量。

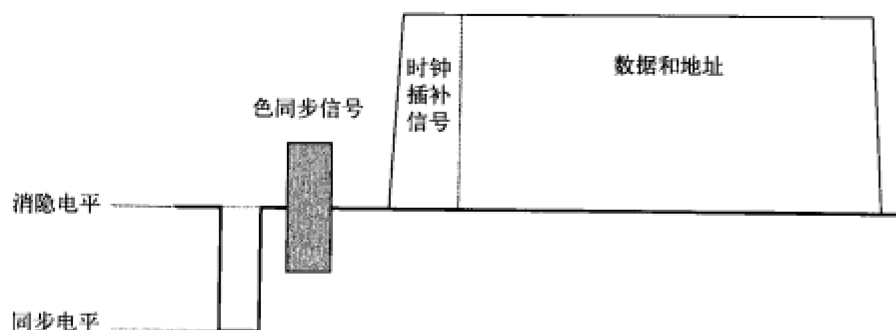


图8-61 图文电视扫描行格式

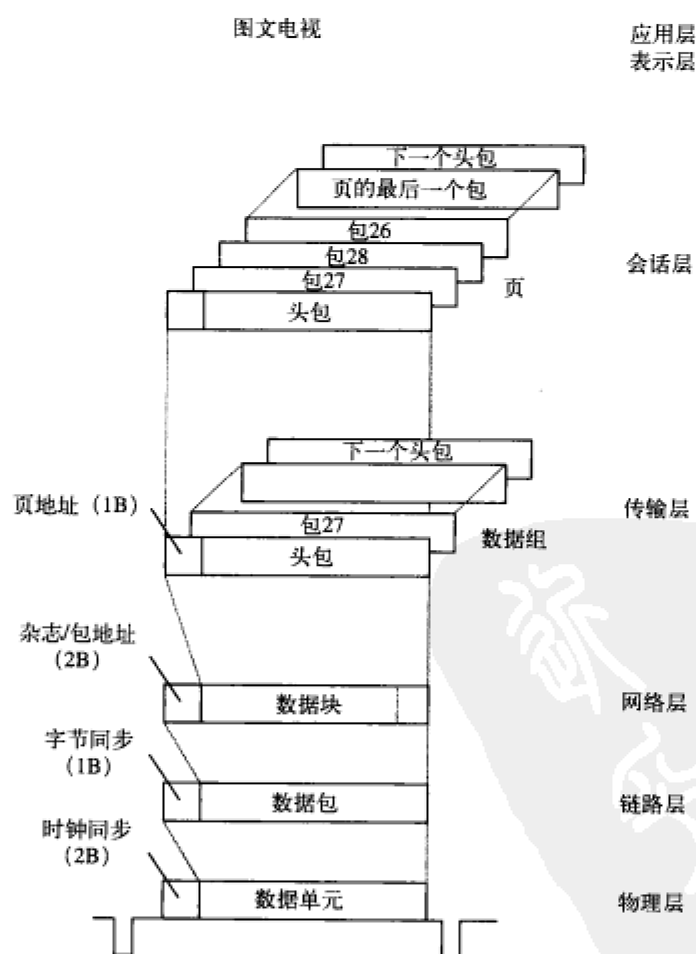


图8-62 图文电视系统B的结构

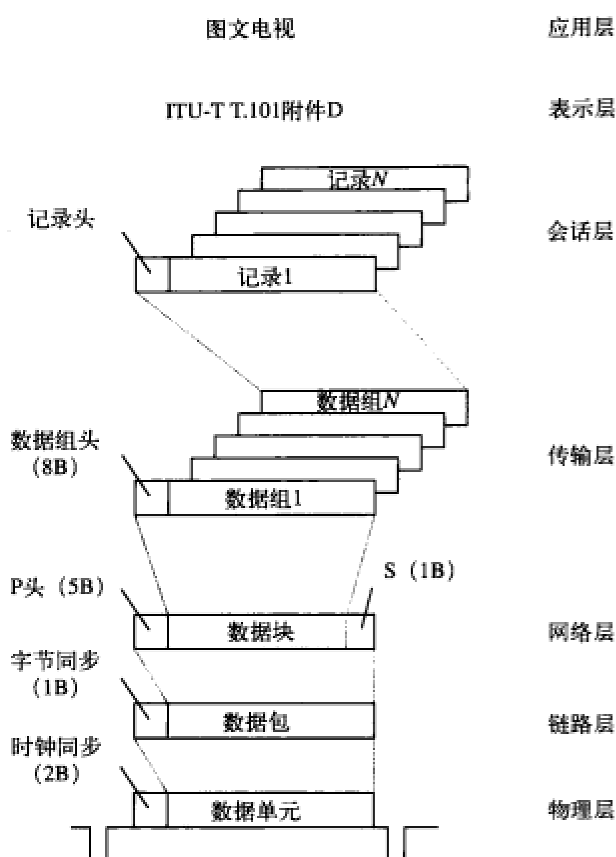


图8-63 图文电视系统C的结构

每个页有24排，每排最多可以有40个字符。字符可以是字母、数字、符号或者简单的图形。图文电视服务中也有控制码，用来选择颜色和其他属性，例如闪烁和双倍高度。

除了传送图文电视信息之外，图文电视协议还可以用来传送其他信息，例如字幕、节目发送控制（program delivery control，PDC）和私有数据。

❑ 字幕

字幕与美国使用的隐藏字幕相似。开放字幕是在传输之前直接插入到图像中的文本。隐藏字幕与图像分开传送。在英国，隐藏字幕的传输使用图文电视第888页。当使用图文电视传送多种语言时，每种语言使用一个单独的页。

375

❑ 节目发送控制（PDC）

节目发送控制（由ETSI ETS 300 231和ITU-R BT.809定义）是一种使用图文电视信息控制VCR录制的制式。可以对VCR进行编程，从而VCR可以查找和录制各种不同类型的节目或者特定的节目。无论节目由于任何原因改变传输时间，都可以继续对节目进行录制。

通过图文电视传送PDC信息的方法有两种：方法A和方法B。

方法A将数据存放在一个可视的图文电视页中，通常在第16扫描行上传送。这种方法被称为视频编程系统（Video Programming System，VPS）。

方法B将数据存放在图文电视信号中的一个隐藏包（包26）中。包26的数据中包含每个节目的数据，包括频道、节目数据和开始时间。

❑ 数据广播

数据广播可以用来向私有接收机传送信息。数据广播的典型应用包括实时财经信息、旅店和旅行社的航空飞行计划、铁路的旅客信息、软件升级等。

❑ 包0~23

一个典型的图文电视页使用24个包，编号为0~23，对应于一张显示页上的24排。包24可以在显示页的底部添加一排状态，用于用户提示。对于每个包，有3位指定杂志地址（1~8），5位指定排地址（0~23）。杂志地址位和排地址位使用汉明（Hamming）码进行误码保护，从而能够纠正单个位的错误。

为了节省带宽，并不是所有的包都发送完整的地址。只有包0［也称为头包（header packet）］中拥有所有的地址信息，例如排、页和杂志的地址数据。包1~28中包含同一个杂志中最近的包0所标识的页的信息。

页的传输以头包开头。随后拥有相同杂志地址的包为页提供额外的数据。这些包可以以任意顺序传送，而且还可以与其他杂志的包交织进行。当接收到该杂志的下一个头包时，才认为这个页传送完毕。

包0的一般格式为：

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
页号	2B
子码	4B
控制码	2B
显示数据	32B

包1~23的一般格式为：

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
显示数据	40B

❑ 包24

这个包为用户提示定义一个额外排。图文电视解码器可以使用包27中的数据响应包24显示排中的用户提示。

376 ~ 378

❑ 包25

这个包定义一个替换标题行（replacement header line）。如果这个包存在，显示40B的数据，而不是包8.30中的频道、页、时间和日期。

❑ 包26

包26由以下内容构成。

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
指示码	1B

13个3字节的数据组，每个数据组包含：7个数据位、6个地址位、5个模式位和6个汉明码位。

包26有15种变体，由指示码定义。包26包含13个数据组，每个数据组指定一个具体的显示位置，以及与该位置相关的数据。

这个包还用于扩展基本字符集的可设置地址范围，从而支持其他语言，例如阿拉伯语、西班牙语、匈牙利语、汉语等。

对于PDC来说，包26中包含每个节目的数据，标识页上节目信息的频道、节目日期、开始时间和光标位置。当用户选择一个节目时，光标位置就链接到合适的包26的预选数据上。然后使用这个数据安排VCR。当节目传送时，节目信息使用包8.30格式2传送。预选数据和包8.30数据之间匹配就可以打开VCR的录制模式。

❑ 包27

包27告诉图文电视解码器如何响应用户对包24的选择。包27最多可以有4个（包27/0~27/3），允许多达24个链接。包27包含以下内容。

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
指示码	1B
链接1（红）	6B
链接2（绿）	6B
链接3（黄）	6B
链接4（青）	6B
链接5（下一页）	6B
链接6（索引）	6B
链接控制数据	1B
页校验数字	2B

每个链接包含：7个数据位、6个地址位、5个模式位和6个汉明码位。

这个包包含将当前页链接到6个页编号（链接）的信息。4个彩色链接对应于远端的4个彩色Fastext页请求键。一般情况下，这4个键对应于显示器底部的使用包24的4个彩色菜单选项。选择一个彩色页请求键就会导致选择相应的链接页。

第五个链接用于指定用户在当前页之后可能希望查看的一个页，例如序列中的下一页。

第六个链接对应于远端的Fastext索引键，并且指定当选择该索引时到达的页地址。

❑ 包28和包29

这两个包用于定义级别2和级别3页，以支持更高分辨率的图像、附加彩色、候选字符集等。379 它们与包26的格式相似。

❑ 包8.30格式1

包8.30（杂志8，包30）不与任何页关联，但是每秒发送一次。这个包也称为电视服务数据包（Television Service Data Packet，TSDP）。它包含通知图文电视解码器传输总体情况和时间的数据。

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
指示码	1B
初始图文电视页	6B
网络ID	2B
相对于UTC的时间偏移	1B
日期（修正儒略日）	3B
UTC时间	3B
电视节目标签	4B

状态显示　　20B

指示码标识传输是位于VBI中还是位于全场中。

初始图文电视页告诉解码器在打开电源时应该捕捉并存储哪个页。通常这是一个索引或者菜单页。

网络标识码标识传输网络。

电视节目标签标识当前节目的节目标签。

状态显示用于显示一个传输状态消息。

- 包8.30格式2

这个格式用于PDC记录器控制，每个流中每秒传送一次。它包含一个指示每个节目开始的节目标签，通常在节目开始前30秒传送，从而允许VCR检测节目，并做好录制节目的准备。

时钟插补信号	2B
帧码	1B
杂志和排地址	2B
指示码	1B
初始图文电视页	6B
标签频道ID	1B
节目控制状态	1B
国家和网络ID	2B
节目ID标签	5B
国家和网络ID	2B
节目类型	2B
状态显示	20B

除了状态显示信息之前的13B信息之外，格式2的内容与格式1的内容相同。

标签频道ID（LCI）标识每个频道可以同时并发传送多达4个PDC流。

节目控制状态（PCS）标识实时状态信息，例如模拟音频传输的类型。

国家和网络ID（CNI）分成两组。第一部分指定国家，第二部分指定网络。

节目ID标签（PIL）指定节目开始的月、日和本地时间。

节目类型（PTY）是一个标识节目受众类型的代码，例如“成人”、“儿童”、“音乐”、“戏剧”等。 [380]

- 包31

包31用于向私有接收机传输数据。

时钟插补信号	2B
帧码	1B
数据通道组	1B
消息位	1B
格式类型	1B
地址长度	1B
地址	0~6B
重播指示器	0~1B
连续指示器	0~1B
数据长度	0~1B
用户数据	28~36B
CRC	2B

8.5.5 AMOL

1. AMOL I

第20、22、283行和第284行用于传送AMOL I信息（AMOL是Automated Measurement Of Lineups的缩写，即广播节目自动测量），如图8-64所示。但是，由于视频编辑、压缩等导致的无意的偏移，AMOL I信息可以位于480i系统的任意VBI行上。1Mbit/s的负载可能随着帧的不同而经常变化。

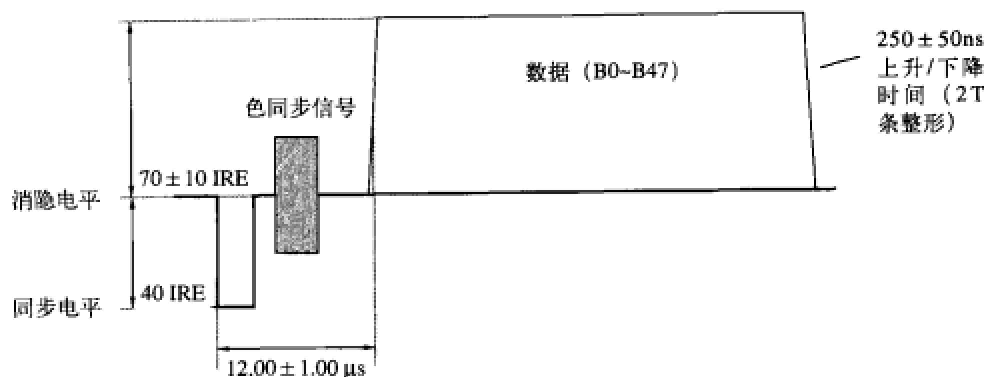


图8-64 525行系统第20、22、283行和第284行AMOL I时序

AMOL I信息每48数据位有1000±100ns宽，最大上升和下降时间为300ns。逻辑1的幅度为55±5 IRE；逻辑0的幅度为0~10 IRE。

2. AMOL II

第20、22、283行和第284行用于传送AMOL II信息，如图8-65所示。但是，由于视频编辑、压缩等导致的无意的偏移，AMOL II信息可以位于480i系统的任意VBI行上。2Mbit/s的负载可能随着帧的不同而经常变化。

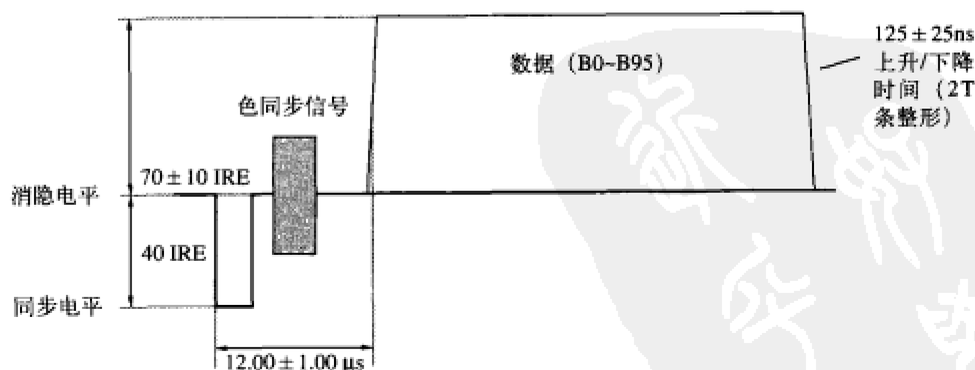


图8-65 525行系统第20、22、283行和第284行AMOL II时序

AMOL II信息每96数据位有500±50ns宽，最大上升和下降时间为150ns。逻辑1的幅度为55±5 IRE；逻辑0的幅度为0~10 IRE。

8.5.6 原始VBI数据

原始或者过采样（oversampled）VBI数据是简单数字化的VBI数据。过采样通常使用2倍的视频采样时钟，例如480i系统使用27MHz，480p系统使用54MHz。使用2倍的视频采样时钟允许在一

个标准的8位BT.656接口上传送原始VBI数据。除了齿脉冲和均衡脉冲期间，VBI数据可以在任何扫描行上提供。

然后原始VBI数据转换为二进制［或者条带（sliced）］数据，并进行处理，然后再传递到复合、S-video和YPbPr模拟视频输出中，从而可以由电视机解码。

在由原始VBI数据向条带VBI数据转换过程中，VBI解码器必须补偿各种DC偏移、幅度变化、重影及时序的变化。

由于噪音和传输误差，还必须使用磁滞（hysteresis）来防止直接打开和关闭VBI解码器。一旦在15个连续帧（典型值）上查找到期望的VBI信号，VBI解码就开始进行。当在45个连续帧（典型值）的合适扫描行上没有找到期望的VBI信号时，VBI解码就应该停止。

8.5.7 条带VBI数据

NTSC/PAL视频解码器通常提供条带或者二进制VBI数据。这样做有一个优点，即因为提供的是二进制数据，而不是过采样的数据，所以需要较低的数据速率。这种方式的主要缺点就是NTSC/PAL视频解码器芯片制造商通过视频接口传送条带VBI数据的技术各不相同。 381

8.5.8 NTSC/PAL解码器注意事项

1. 隐藏字幕

除了清除显示的字幕和文本命令外，还有5种其他事件通常可以强制清除显示。

- 字幕显示模式的改变，例如由CC1改变为T1。
- 视频锁的丢失（诸如在频道改变时），强制清除显示。当前有效显示模式不改变。例如，如果在视频锁丢失之前选择的模式是CC1，那么在视频锁丢失之后模式仍然是CC1。
- 自动消隐激活。如果没有在15个连续帧（典型值）上检测到字幕信号，或者在45个连续帧（典型值）上没有接收到所选择频道的新数据，那么显示存储器就被清除。一旦在15个连续帧（典型值）上检测到字幕信号，或者接收到新数据，就显示字幕信号。
- 清除命令（例如来自远端控制器）强制清除显示。
- 禁用字幕解码也会强制清除显示。

2. 宽屏信令

解码器必须能够处理不同的WSS输入，包括：

- 复合、S-video或者Y（YPbPr）视频上的PAL或NTSC制式WSS信号；
- SCART模拟输入（DC偏移指示器）；
- S-video模拟输入（DC偏移指示器）。

除了基于WSS数据自动处理视频信号以适应4：3或者16：9显示之外，解码器还应该支持手动设置4：3或者16：9显示，从而可以适应个人的偏好，按个人的愿望设置特定的显示模式。软件使用该宽高比信息、用户偏好和显示格式帮助正确处理节目的显示。

8.5.9 重影消除

因为一些服务（例如图文电视）的高速度传输，所以要求提供重影消除（移除信号中不希望的反射）。大于100ns或者－12dB的重影会破坏图文电视数据。大于－3dB的重影在硬件或者软件中就难以有效移除，而小于－12dB的重影不需要移除。VBI数据的重影消除不像有效视频的重影消除那样复杂。

然而，GCR（Ghost Cancellation Reference，重影消除基准）信号通常不会提供。因此，重影

消除算法必须使用其他可用信号（例如齿脉冲和均衡脉冲）确定重影量。

NTSC制式GCR信号由ATSC A/49和ITU-R BT.1124指定。如果这个信号存在，那么这个信号位于第19行和第282行上。GCR允许检测－3～＋45μs的重影，后面是一个8场的序列。

382～383 PAL制式GCR信号由BT.1124和ETSI ETS 300 732指定。如果这个信号存在，那么这个信号位于第318行上。GCR允许检测－3～＋45μs的重影，后面是一个4帧的序列。

8.6 增强电视编程

增强电视编程（enhanced television programming）标准（SMPTE 363M）用于创建和发送增强和交互式的节目。增强内容可以通过多种媒介发送，这些媒介包括使用陆地、有线和卫星网络的模拟和数字电视广播。在定义如何创建增强内容的过程中，规范定义了接收机的最低功能。为了减少创建新规范，它利用了因特网技术，例如HTML和Java-script。这样做的好处就是早已拥有数百万的潜在内容页面，而且还可以利用现有的网络工作。

规范规定了接收机最低支持HTML 4.0、Javascript 1.1和层叠样式表（Cascading Style Sheet）。规范还定义了额外的能力，例如Java和VRML，这些是可选的。这样可以确保内容可以被最多的观众观看。

为了支持日益增长的功能，HTML中添加了一个新的“tv：”属性。该属性允许在内容中插入电视节目，而且还可以在HTML文档中存放正常图像的任意位置上使用。创建一个在显示器的任意位置上显示当前电视频道的增强内容页与在HTML文档中插入一个图像一样容易。

该规范还定义了接收机如何获取内容，以及如何通知接收机存在增强内容。后面的一项任务使用触发器（trigger）完成。

8.6.1 触发器

触发器提醒接收机注意内容增强，而且包含有节目增强方面的信息。其他方面，触发器包含一个统一资源定位器（Universal Resource Locator，URL），定义增加内容的位置。内容可能就位于本地（例如通过网络传送时缓冲到本地硬盘上），也有可能位于因特网或者另一个网络上。

触发器还可以包含适合于人类阅读的内容描述。例如，它可以包含这样的描述：Press ORDER to order this product.。该描述可以显示给观众。触发器还可以包含截止日期信息，向观众提示增强服务还可以提供多久。

最后，触发器可以包含触发关联HTML页中Javascript执行的脚本，支持增强内容与视频信号的同步，以及动态屏幕数据的更新。

触发器的处理由SMPTE 363M定义，与传递它们的方法无关。

8.6.2 传输

除了定义内容如何显示以及如何通知接收机新内容之外，增强电视节目规范还定义了如何传递内容。因为接收机可能没有因特网连接，所以规范描述了两种传递内容的模型。这两个模型都称为传输（transport），这两个传输被称为传输类型A和传输类型B。

如果接收机拥有一个到因特网的回返通道（back-channel），或者返回路径，那么传输类型A将广播触发器，内容将通过因特网推出去。

384 如果接收机没有因特网连接，那么传输类型B通过广播媒介提供触发器和内容的传递。声明通过网络发送出去，从而将触发器和内容流关联。声明描述内容，而且还可以包含有关带宽、存储需求及语言的信息。

8.6.3 传递协议

对于传统的双向因特网通信来说，超文本传输通信协议（Hypertext Transfer Protocol，HTTP）定义了如何在应用层传送数据。对于单向的广播来说，这种情况没有双向的连接，因此SMPTE 364M为数据传递定义了一个单向的应用层协议：单向超文本传输通信协议（Uni-directional Hyper Text Transfer Protocol，UHTTP）。

像HTTP一样，UHTTP使用传统的URL命名方案引用内容。内容可以使用标准的"http："和"ftp："命名方案引用增强页，还可以使用一个"lid："或者本地标识符、URL，引用本地存在的内容（例如接收机的硬盘上），而不是因特网或者其他网络上的内容。

8.6.4 绑定

数据如何通过一个特定的网络进行传递称之为"绑定"。NTSC和PAL制式都定义了绑定。

1. NTSC制式绑定

传输类型A触发器在CEA-608字幕信号的数据通道2上广播。

传输类型B绑定也包含一种通过VBI传递IP多播包的机制，这也可以称为VBI期IP传送（IP/VBI）。在最底层次，电视信号在VBI期间传送NABTS（North American Basic Teletext Standard，北美基本图文电视标准）包。这些NABTS包恢复以后，形成一个连续的数据流（封装在一个与SLIP相似的协议内），解帧后生成IP包。

2. PAL制式绑定

两种传输类型都基于PAL制式的VBI行中通过图文电视包30或31携带的IP多播包。

传输类型A触发器通过UDP/IP多播包携带，传递到地址224.0.23.13和端口2670上。

传输类型B（在SMPTE 357M中描述）在一个单独的UDP/IP多播包中携带一个单独的触发器，传递到增强电视节目的SDP声明中定义的地址和端口上。为了提供快速的同步，所定义的触发器协议非常轻便。

参考文献

1. Advanced Television Enhancement Forum, *Enhanced Content Specification*, 1999.
2. ATSC A/49, 13 May 1993, *Ghost Cancelling Reference Signal for NTSC.*
3. *BBC Technical Requirements for Digital Television Services*, Version 1.0, February 3, 1999, BBC Broadcast.
4. CEA-608, August 2005, *Line 21 Data Service.*
5. EIA-189-A, July 1976, *Encoded Color Bar Signal.*
6. EIA-516, May 1988, *North American Basic Teletext Specification (NABTS).*
7. EIA-J CPR-1204, *Transfer Method of Video ID Information using Vertical Blanking Interval (525-line System)*, March 1997.
8. ETSI EN 300 163, *Television Systems: NICAM 728: Transmission of Two Channel Digital Sound with Terrestrial Television Systems B, G, H, I, K1, and L*, March 1998.
9. ETSI EN 300 231, *Television Systems: Specification of the Domestic Video Programme Delivery Control System (PDC)*, April 2003.
10. ETSI EN 300 294, *Television Systems: 625-line Television Widescreen Signalling (WSS)*, April 2003.
11. ETSI EN 300 706, *Enhanced Teletext Specification*, April 2003.
12. ETSI EN 300 708, *Television Systems: Data Transmission within Teletext*, April 2003.
13. ETSI ETS 300 731, *Television Systems: Enhanced 625-Line Phased Alternate Line (PAL) Television: PALplus*, March 1997.
14. ETSI ETS 300 732, *Television Systems: Enhanced 625-Line PAL/SECAM Television; Ghost Cancellation Reference (GCR) Signals*, January 1997.
15. Faroudja, Yves Charles, *NTSC and Beyond, IEEE Transactions on Consumer Electronics*, Vol. 34, No. 1, February 1988.
16. IEC 61880, 1998-1, *Video Systems (525/60)—Video and Accompanied Data Using the Vertical Blanking Interval—Analog*

385

Interface.
17. ITU-R BS.707-3, 1998, *Transmission of Multisound in Terrestrial Television Systems PAL B, G, H, and I and SECAM D, K, K1, and L.*
18. ITU-R BT.470-6, 1998, *Conventional Television Systems.*
19. ITU-R BT.471-1, 1986, *Nomenclature and Description of Colour Bar Signals.*
20. ITU-R BT.472-3, 1990, *Video Frequency Characteristics of a Television System to Be Used for the International Exchange of Programmes Between Countries that Have Adopted 625-Line Colour or Monochrome Systems.*
21. ITU-R BT.473-5, 1990, *Insertion of Test Signals in the Field-Blanking Interval of Monochrome and Colour Television Signals.*
22. ITU-R BT.569-2, 1986, *Definition of Parameters for Simplified Automatic Measurement of Television Insertion Test Signals.*
23. ITU-R BT.653-3, 1998, *Teletext Systems.*
24. ITU-R BT.809, 1992, *Programme Delivery Control (PDC) System for Video Recording.*
25. ITU-R BT.1118, 1994, *Enhanced Compatible Widescreen Television Based on Conventional Television Systems.*
26. ITU-R BT.1119-2, 1998, *Wide-Screen Signalling for Broadcasting.*
27. ITU-R BT.1124, 1994, *Reference Signals for Ghost Cancelling in Analogue Television Systems.*
28. ITU-R BT.1197-1, 1998, *Enhanced Wide-Screen PAL TV Transmission System (the PALplus System).*
29. ITU-R BT.1298, 1997, *Enhanced Wide-Screen NTSC TV Transmission System.*
30. *Multichannel TV Sound System BTSC System Recommended Practices*, EIA Television Systems Bulletin No. 5 (TVSB5).
31. *NTSC Video Measurements*, Tektronix, Inc., 1997.
32. SMPTE 12M-1999, *Television, Audio and Film—Time and Control Code.*
33. SMPTE 170M-2004, *Television—Composite Analog Video Signal—NTSC for Studio Applications.*
34. SMPTE 262M-1995, *Television, Audio and Film—Binary Groups of Time and Control Codes—Storage and Transmission of Data.*
35. SMPTE 309M-1999, *Television—Transmission of Date and Time Zone Information in Binary Groups of Time and Control Code.*
36. SMPTE 357M-2002, *Television—Declarative Data Essence —Internet Protocol Multicast Encapsulation.*
386
37. SMPTE 361M-2002, *Television—NTSC IP and Trigger Binding to VBI.*
38. SMPTE 363M-2002, *Television—Declarative Data Essence —Content Level 1.*
39. SMPTE 364M-2001, *Declarative Data Essence — Unidirectional Hypertext Transport Protocol.*
40. SMPTE RP-164-1996, *Location of Vertical Interval Time Code.*
41. SMPTE RP-186-1995, *Video Index Information Coding for 525- and 625-Line Television Systems.*
42. SMPTE RP-201-1999, *Encoding Film Transfer Information Using Vertical Interval Time Code.*
43. *Specification of Television Standards for 625-Line System-I Transmissions*, 1971, Independent Television Authority (ITA) and British Broadcasting Corporation (BBC).
44. *Television Measurements, NTSC Systems*, Tektronix, Inc., 1998.
387
45. *Television Measurements, PAL Systems*, Tektronix, Inc., 1990.

第9章 NTSC和PAL数字编码和解码

尽管准确地说，NTSC和PAL复合彩色视频格式并不是数字视频，但是它们是目前最常见的视频格式。尽管视频信号本身是模拟的，但是它们几乎完全采用数字式编码和解码。

模拟的NTSC和PAL编码器和解码器已经面世好长一段时间了。但是，它们使用起来非常困难，需要进行调整，并且视频质量受限。通过使用数字技术来实现NTSC和PAL编码和解码有许多优点，例如使用方便、最小化的模拟调整以及出色的视频质量。

除了复合视频，家用和专业视频设备还支持S-Video，所以也应该实现S-Video的数字编码和解码。S-Video使用单独的亮度（Y）和色度（C）模拟视频信号，因此通过消除Y/C的分离过程，可以实现高质量的视频。

本章讨论支持复合和S-Video的（M）NTSC和（B、D、G、H、I、N_C）PAL视频信号的数字编码器（图9-1）和解码器（图9-21）的设计。只需进行一些细小的修改，就可以实现（M）和（N）PAL视频信号的数字编码器和解码器。

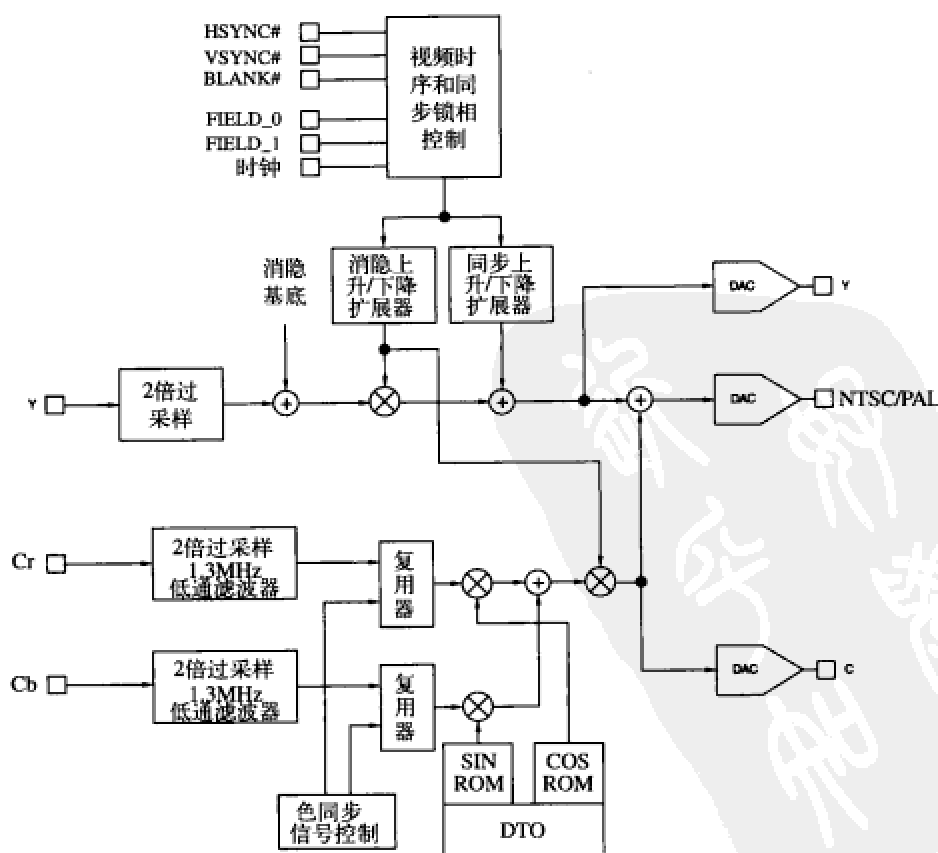


图9-1 典型NTSC/PAL数字编码器实现

NTSC编码器和解码器通常基于YCbCr、YUV或者YIQ颜色空间。PAL编码器和解码器通常基

388 于YCbCr或者YUV颜色空间。

9.1　NTSC和PAL编码

YCbCr输入数据的标称范围为：Y为16~235，Cb和Cr为16~240。RGB输入数据的标称范围为0~255；专业视频应用使用的标称范围为16~235。

当YCbCr的值位于这些范围之外时，将致使一些颜色组合溢出标准的YIQ或YUV范围，该问题可以通过下述3种方式解决，按优先顺序排列：（1）允许视频信号使用扩展的YIQ或YUV范围生成；（2）限制色饱和度确保生成有效的视频信号；（3）将YIQ或YUV电平削波到有效范围。

4：1：1、4：2：0或4：2：2的YCbCr数据在转换成YIQ或者YUV数据之前，必须先转换成4：4：4的YCbCr数据，否则色度低通滤波器无法正确地执行插值。

表9-1列举了一些常用的采样率和分辨率。

表9-1　常用的NTSC/PAL采样率和分辨率

视频标准	采样时钟频率	应　用	有效分辨率	总分辨率	场频（/秒）
（M）NTSC、（M）PAL	9MHz	SVCD	480×480i	572×525i	59.94隔行
	13.5MHz	BT.601	720[1]×480i	858×525i	
		MPEG-2	704×480i		
		DV	720×480i		
	12.27MHz	方像素	640×480i	780×525i	
（B、D、G、H、I、N、N_C）PAL	9MHz	SVCD	480×576i	576×625i	50隔行
	14.75MHz	方像素	768×576i	944×625i	
	13.5MHz	BT.601	720[2]×576i	864×625i	
		MPEG-2	704×576i		
		DV	720×576i		

注：1. 10%消隐点之间拥有716个典型的实有效样本。
　　2. 50%消隐点之间拥有702个典型的实有效样本。

9.1.1　2倍过采样

2倍过采样生成8：8：8的YCbCr数据或者RGB数据，简化模拟输出滤波器。过采样器（oversampler）还是一个将数据由8位转换为10位的方便位置，从而提高视频的质量。

9.1.2　颜色空间转换

对于10位的视频，选择电平为：白色＝800，同步＝16，并且已知同步－白色幅度为1V，因389 此将数字模拟转换器（D/A Converter，DAC）的完整量程输出设置为1.305V。

1. （M）NTSC、（M、N）PAL

因为（M）NTSC、（M、N）PAL的消隐基底为7.5 IRE，同步幅度为40 IRE，所以导出的颜色空间转换方程要生成0.660V的有效视频。

❑ YUV颜色空间处理

现代的编码器设计都基于YUV颜色空间，对于这些编码器，由YCbCr生成YUV的方程如下：

$Y = 0.591(Y_{601} - 64)$

$U = 0.504(Cb - 512)$

$V = 0.711(Cr - 512)$

由R'G'B'生成YUV的方程如下：

$Y = 0.151R' + 0.297G' + 0.058B'$

$U = -0.074R' - 0.147G' + 0.221B'$

$V = 0.312R' - 0.261G' - 0.051B'$

对于为RGB使用10位标称范围64~940的专业视频应用来说，由R'G'B'生成YUV的方程如下：

$Y = 0.177(R' - 64) + 0.347(G' - 64) + 0.067(B' - 64)$

$U = -0.087(R' - 64) - 0.171(G' - 64) + 0.258(B' - 64)$

$V = 0.364(R' - 64) - 0.305(G' - 64) - 0.059(B' - 64)$

Y的标称范围为0~518，U的标称范围为0~±226，V的标称范围为0~±319。为了允许测试信号、抠像信息及实时视频以最小的失真通过编码器，Y还应该支持负值。

❑ YIQ颜色空间处理

对于基于YIQ颜色空间的较老的NTSC编码器设计，由YCbCr生成YIQ的方程如下：

$Y = 0.591(Y_{601} - 64)$

$I = 0.596(Cr - 512) - 0.274(Cb - 512)$

$Q = 0.387(Cr - 512) + 0.423(Cb - 512)$

由R'G'B'生成YIQ的方程如下：

$Y = 0.151R' + 0.297G' + 0.058B'$

$I = 0.302R' - 0.139G' - 0.163B'$

$Q = 0.107R' - 0.265G' + 0.158B'$

对于为R'G'B'使用10位标称范围64~940的专业视频应用来说，由R'G'B'生成YIQ的方程如下：

$Y = 0.177(R' - 64) + 0.347(G' - 64) + 0.067(B' - 64)$

$I = 0.352(R' - 64) - 0.162(G' - 64) - 0.190(B' - 64)$

$Q = 0.125(R' - 64) - 0.309(G' - 64) + 0.184(B' - 64)$

Y的标称范围为0~518，I的标称范围为0~±309，Q的标称范围为0~±271。为了允许测试信号、抠像信息及实时视频以最小的失真通过编码器，Y还应该支持负值。

❑ YCbCr颜色空间处理

如果编码器的设计基于YUV颜色空间，那么通过在调制过程中缩放正弦值和余弦值或者缩放色差低通滤波器系数，可以避免将Cb和Cr转换为U和V。这样做的优点在于可以减少数据路径处理。 390~391

2. NTSC-J

因为在日本使用的（M）NTSC采用一个0 IRE的消隐基底，所以导出的颜色空间转换方程要生成0.714V的有效视频。

❑ YUV颜色空间处理

由YCbCr生成YUV的方程如下：

$Y = 0.639(Y_{601} - 64)$

$U = 0.545(Cb - 512)$

$V = 0.769(Cr - 512)$

由R'G'B'生成YUV的方程如下：

$Y = 0.164R' + 0.321G' + 0.062B'$

$U = -0.080R' - 0.159G' + 0.239B'$

$V = 0.337R' - 0.282G' - 0.055B'$

对于为R'G'B'使用10位标称范围64~940的专业视频应用来说，由R'G'B'生成YUV的方程如下：

$Y = 0.191(R' - 64) + 0.375(G' - 64) + 0.073(B' - 64)$

$U = -0.094(R'-64) - 0.185(G'-64) + 0.279(B'-64)$

$V = 0.393(R'-64) - 0.329(G'-64) - 0.064(B'-64)$

Y的标称范围为0~560，U的标称范围为0~±244，V的标称范围为0~±344。为了允许测试信号、抠像信息及实时视频以最小的失真通过编码器，Y还应该支持负值。

❑ YIQ颜色空间处理

对于基于YIQ颜色空间的较老的编码器设计，由YCbCr生成YIQ的方程如下：

$Y = 0.639(Y_{601} - 64)$

$I = 0.645(Cr - 512) - 0.297(Cb - 512)$

$Q = 0.419(Cr - 512) + 0.457(Cb - 512)$

由R′G′B′生成YIQ的方程如下：

$Y = 0.164R' + 0.321G' + 0.062B'$

$I = 0.326R' - 0.150G' - 0.176B'$

$Q = 0.116R' - 0.286G' + 0.170B'$

对于为R′G′B′使用10位标称范围64~940的专业视频应用来说，由R′G′B′生成YIQ的方程如下：

$Y = 0.191(R'-64) + 0.375(G'-64) + 0.073(B'-64)$

$I = 0.381(R'-64) - 0.176(G'-64) - 0.205(B'-64)$

$Q = 0.135(R'-64) - 0.334(G'-64) + 0.199(B'-64)$

Y的标称范围为0~560，I的标称范围为0~±334，Q的标称范围为0~±293。为了允许测试信号、抠像信息及实时视频以最小的失真通过编码器，Y还应该支持负值。

❑ YCbCr颜色空间处理

如果编码器的设计基于YUV颜色空间，那么通过在调制过程中缩放正弦值和余弦值或者缩放色差低通滤波器系数，可以避免将Cb和Cr转换为U和V。这样做的优点在于可以减少数据路径处理。

392

3.（B、D、G、H、I、N_C）PAL

因为这些PAL标准的消隐基底为0 IRE，同步幅度为43 IRE，所以导出的颜色空间转换方程要生成0.7V的有效视频。

❑ YUV颜色空间处理

由YCbCr生成YUV的方程如下：

$Y = 0.625(Y_{601} - 64)$

$U = 0.533(Cb - 512)$

$V = 0.752(Cr - 512)$

由R′G′B′生成YUV的方程如下：

$Y = 0.160R' + 0.314G' + 0.061B'$

$U = -0.079R' - 0.155G' + 0.234B'$

$V = 0.329R' - 0.275G' - 0.054B'$

对于为R′G′B′使用10位标称范围64~940的专业视频应用来说，由R′G′B′生成YUV的方程如下：

$Y = 0.187(R'-64) + 0.367(G'-64) + 0.071(B'-64)$

$U = -0.092(R'-64) - 0.181(G'-64) + 0.273(B'-64)$

$V = 0.385(R'-64) - 0.322(G'-64) - 0.063(B'-64)$

Y的标称范围为0~548，U的标称范围为0~±239，V的标称范围为0~±337。为了允许测试信号、抠像信息及实时视频以最小的失真通过编码器，Y还应该支持负值。

❑ YCbCr颜色空间处理

如果编码器的设计基于YUV颜色空间，那么通过在调制过程中缩放正弦值和余弦值或者缩放

色差低通滤波器系数，可以避免将Cb和Cr转换为U和V。这样做的优点在于可以减少数据路径处理。

9.1.3 亮度（Y）处理

为了移除2倍过采样过程中生成的高频分量，必须进行6MHz的低通滤波。

为了从亮度信息中移除彩色副载波频率，还可以使用一个可选的陷波滤波器。这样可以改进使用简单Y/C分离解码器的解码视频质量。当生成S-Video、RGB或者YPbPr视频信号时，应该禁用陷波滤波器。

接下来，在有效视频中添加消隐基底以及消隐和同步信息。

1. （M）NTSC、（M、N）PAL

因为（M）NTSC、（M、N）PAL的消隐基底为7.5 IRE，所以在有效视频中将亮度数据加上一个值42。在消隐过程中添加的值为0。

在添加了消隐基底之后，亮度数据通过一个拥有凸余弦分布的消隐信号钳制，以减缓视频信号起始点和结束点的摆率（slew rate）。NTSC制式消隐上升和下降时间的一般值为140±20ns，PAL制式的一般值为300±100ns。

在执行了消隐处理之后，数字复合同步信息添加到亮度数据上。有同步时所指定的值为16，没有同步时所指定的值为240。对同步上升和下降时间的处理应该生成一个凸余弦分布（位于16至240之间），以减缓同步信号的摆率。尽管编码器应该生成130ns或240ns的同步边，以补偿模拟输出滤波器对同步边的减缓，但是NTSC制式同步上升和下降时间的一般值为140±20ns，PAL制式的一般值为250±50ns。 393

此时，我们拥有了数字的亮度及同步和消隐信息，如表9-2所示。

表9-2 10位数字亮度值

视频电平	（M）NTSC	NTSC-J	（B、D、G、H、I、N_C）PAL	（M、N）PAL
白色电平	800	800	800	800
黑色电平	282	240	252	282
消隐电平	240	240	252	240
同步电平	16	16	16	16

2. NTSC-J

当生成NTSC-J视频时，消隐基底为0 IRE。因此，有效视频期间的亮度数据没有添加消隐基底。否则，处理过程就和（M）NTSC一样。

3. （B、D、G、H、I、N_C）PAL

当生成（B、D、G、H、I、N_C）PAL视频时，消隐基底为0 IRE。因此，有效视频期间的亮度数据没有添加消隐基底。

消隐信息使用和（M）NTSC相同的技术处理。但是，（B、D、G、H、I、N_C）PAL的典型消隐上升和下降时间为300±100ns。

复合同步信息使用和（M）NTSC相同的技术添加，只是当提供同步时使用的值为15，不提供同步时使用的值为252。尽管编码器应该生成大约240ns的同步边，以补偿模拟输出滤波器对同步边的减缓，但是同步上升和下降时间的典型值为250±50ns。

此时，我们拥有了数字的亮度及同步和消隐信息，如表9-2所示。

4. 模拟亮度（Y）生成

数字亮度数据可以驱动一个生成0~1.305V输出的10位DAC，以生成S-Video（Y/C）接口的Y

视频信号。

图9-2和图9-3示范了75%彩条的亮度视频波形。亮度电平上的数字表示满量程输出值为1.305V的10位DAC的数据值。连接器上的视频信号应该有一个75Ω的源阻抗。

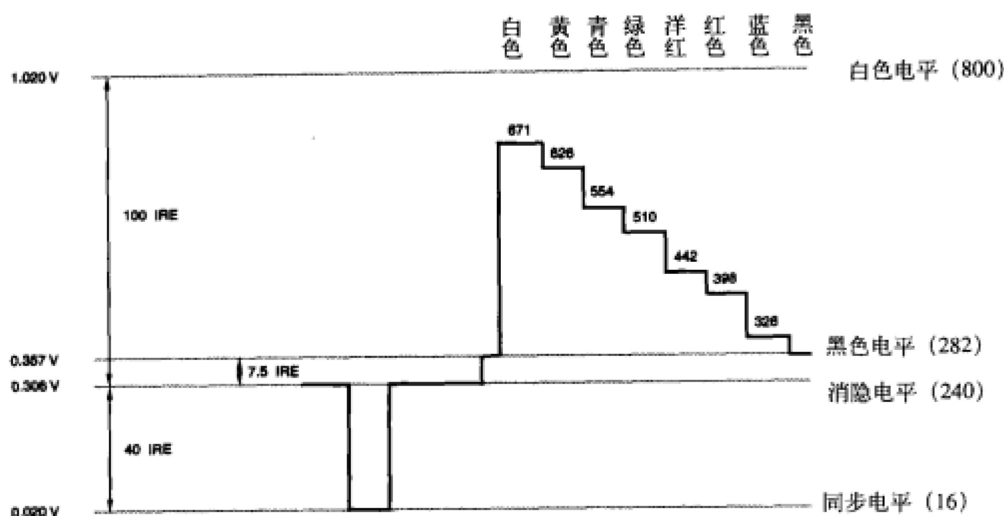


图9-2 75%彩条的（M）NTSC亮度（Y）视频信号。图中指示的亮度电平值为10位

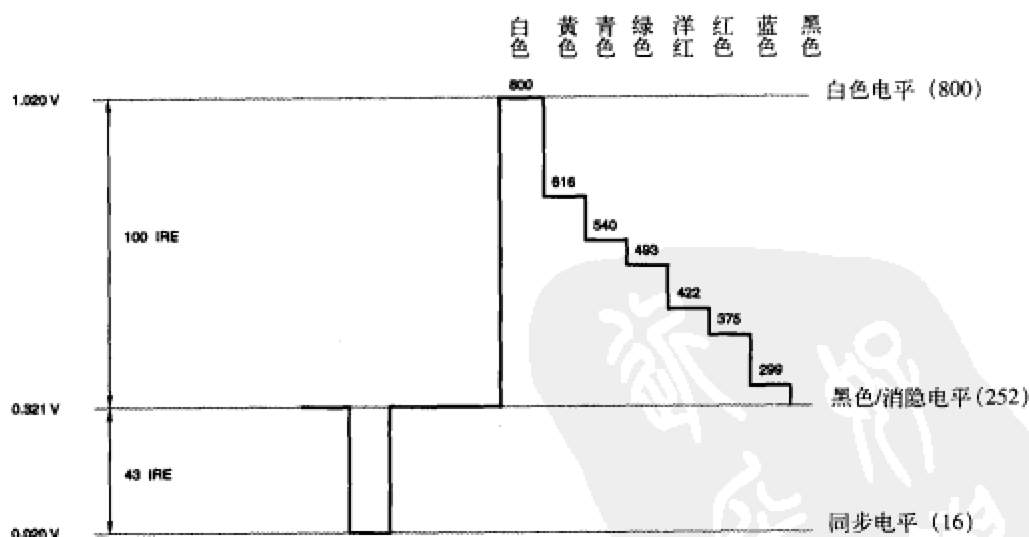


图9-3 75%彩条的（B、D、G、H、I）PAL亮度（Y）视频信号。图中指示的亮度电平值为10位

由于DAC的采样/保持动作引入了一个$(\sin x)/x$特性，所以视频数据可以通过一个$[(\sin x)/x]^{-1}$滤波器的数字滤波进行补偿。另外，因为通常在DAC之后提供一个模拟低通滤波器，所以校正可以在模拟滤波器中进行。

394

作为一个选项，提供在驱动DAC之前延迟数字Y信息一个可编程的时钟周期数量的能力是有用的。如果模拟亮度视频在DAC之后低通滤波，而且模拟色度视频在它的DAC之后进行带通滤波，那么色度视频路径的延迟（典型值大约为400ns）要比亮度视频路径的延迟长。通过调整Y数据的延迟，模拟亮度和色度视频在滤波之后将更加对准，从而简化模拟设计。

9.1.4 色差处理

1. 低通滤波

色差信号（CbCr、UV或者IQ）应该使用一个高斯滤波器低通滤波。这种类型的滤波器能够使振铃（ringing）和过冲（overshoot）最小，避免在陡峭边沿上生成可视的伪影。

如果在视频编辑应用中使用编码器，那么滤波器在通带中的波动应该不大于±0.1dB。这样就可以使由滤波器引起的增益和损失伪影的累积最小，特别是当同时有多路编码和解码时。在编码的最后，可以使用高斯滤波器。

❑ YCbCr和YUV颜色空间

Cb和Cr，或者U和V，都进行1.3MHz的低通滤波。滤波器的典型特性为在1.3MHz处衰减小于2dB，在3.6MHz处衰减大于20dB。滤波器的特性如图9-4所示。

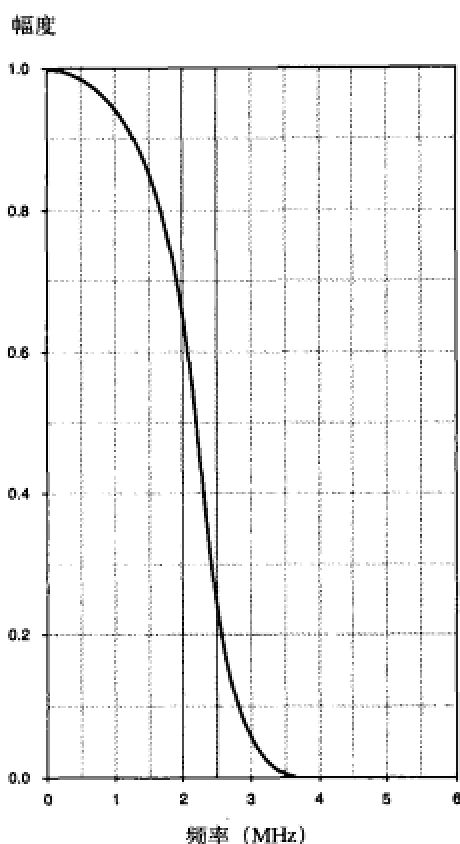


图9-4 典型1.3MHz低通数字滤波器特性

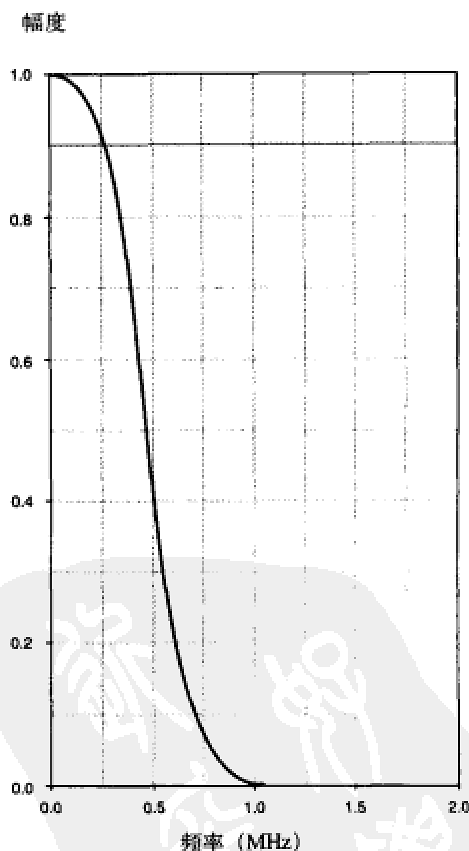


图9-5 典型0.6MHz低通数字滤波器特性

❑ YIQ颜色空间

Q进行0.6MHz的低通滤波。滤波器的典型特性为在0.4MHz处衰减小于2dB，在0.5MHz处衰减小于6dB，在0.6MHz处衰减大于6dB。滤波器的特性如图9-5所示。

I的典型滤波器特性与U和V的相同。

❑ 滤波器注意事项

调制过程以频谱形式在图9-6~图9-9中示范。如果调制过程是模拟的，那么调制过程的频谱都是相同的，只是在采样率的谐频处重复。

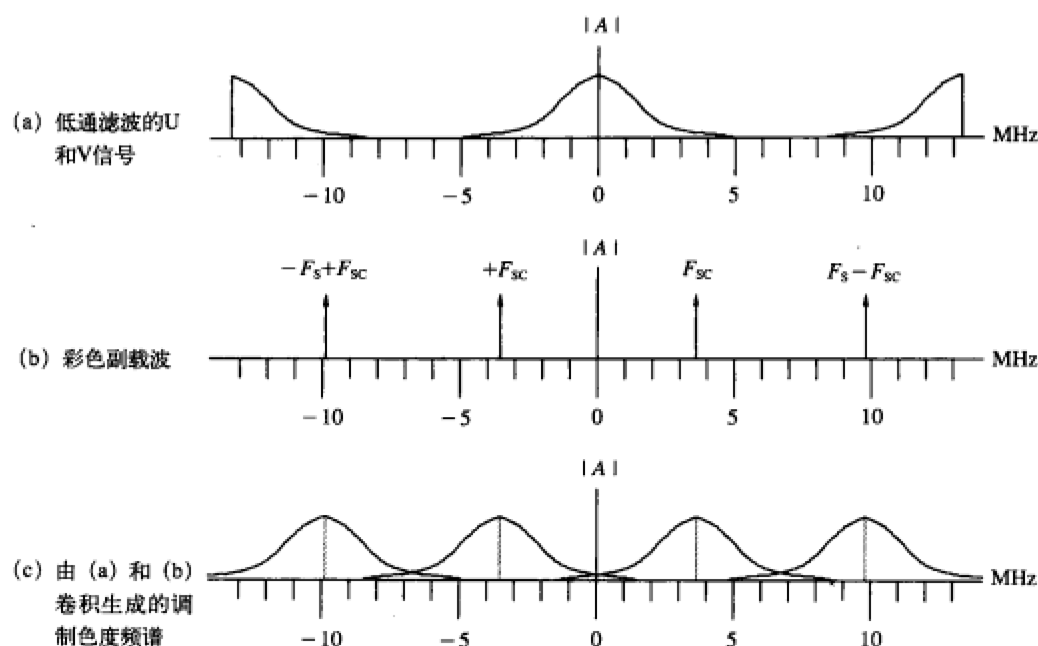


图9-6 NTSC数字色度调制的频谱（$F_S=13.5\text{MHz}$、$F_{SC}=3.58\text{MHz}$）

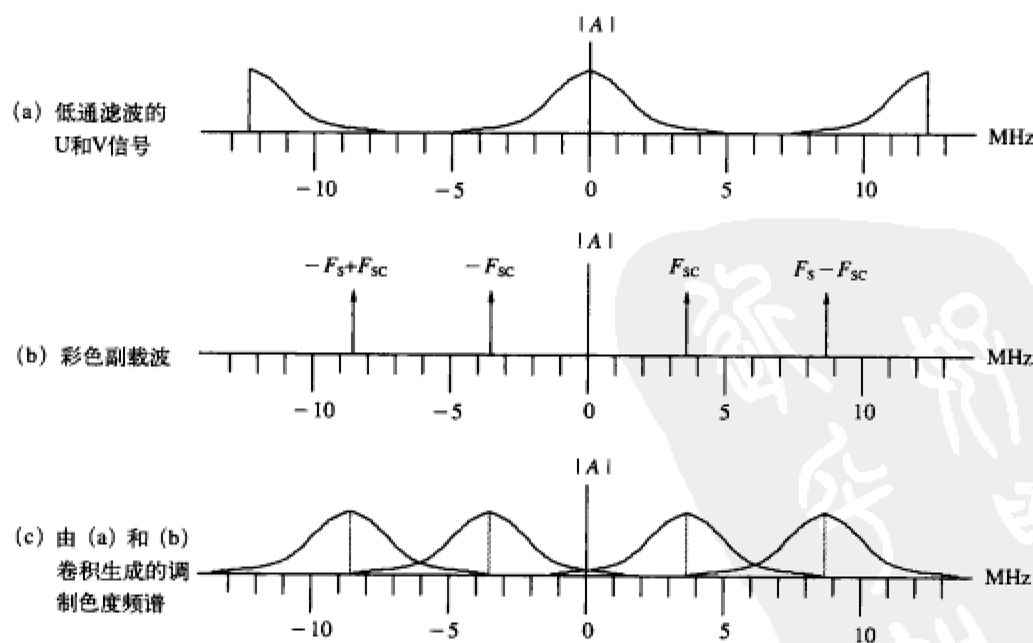


图9-7 NTSC数字色度调制的频谱（$F_S=12.27\text{MHz}$、$F_{SC}=3.58\text{MHz}$）

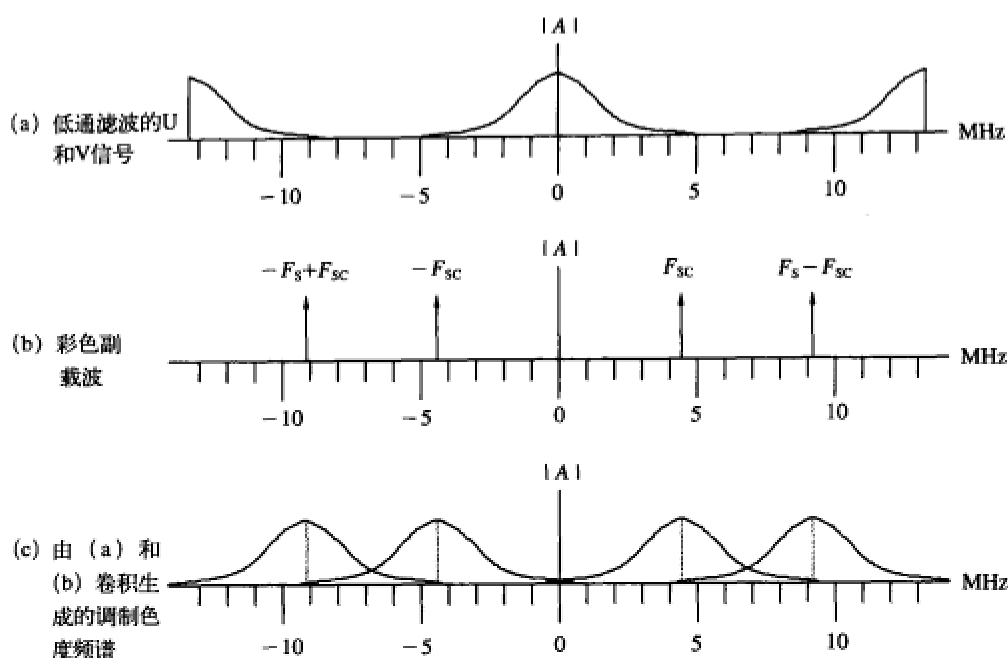


图9-8 PAL数字色度调制的频谱（$F_S=13.5\text{MHz}$、$F_{SC}=4.43\text{MHz}$）

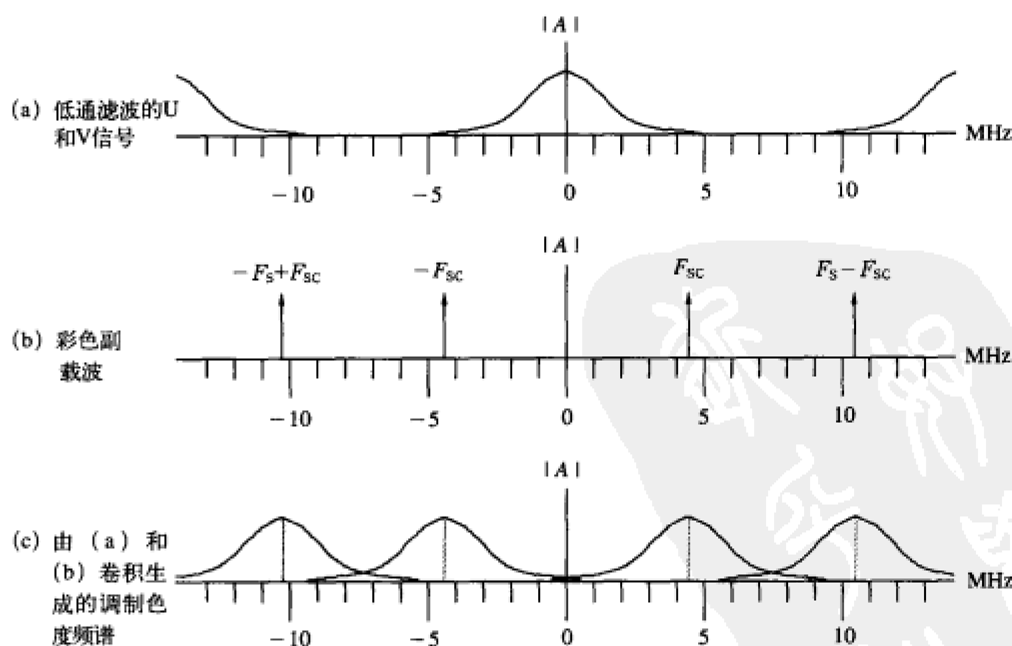


图9-9 PAL数字色度调制的频谱（$F_S=14.75\text{MHz}$、$F_{SC}=4.43\text{MHz}$）

通过使用宽带（1.3MHz）滤波器，调制色度频谱在0频率区域附近重叠，导致产生混叠。另外，在副载波频率之上可能还会存在相当的混叠。出于这些原因，使用较窄带的低通滤波器（0.6MHz）可能更合适。

宽带高斯滤波器能够通过降低彩色对象边缘的伪影，确保与黑白显示器最佳兼容。带宽较窄、

较锐利的低通滤波器将在这些边上加重副载波信号，导致振铃。如果可以忽略黑白兼容性，那么使用较窄带滤波器的好处就是将色度的传播归并到低频亮度之中（致使低频亮度串扰），而在解码器中抑制低频亮度串扰非常困难。

另外，尽管编码器可能维持一个宽的色度带宽，但是解码器中色差信号的带宽通常非常窄。在解码器中，色度上边带的损失（由于将视频信号低通滤波为4.2~5.5MHz）致使产生振铃以及彩色转换上的色差串扰。解码器色度带宽的增加将致使串色成比例的递增。

395～396

2. 色度（C）调制

❑（M）NTSC、NTSC-J

在有效视频期间，CbCr、UV或者IQ数据调制正弦和余弦副载波，如图9-1所示，结果生成数字色度（C）数据。对于该设计，11位基准副载波相位（参见图9-17）和色同步信号相位相同（180°）。

对于YUV和YCbCr处理，在有效视频时间期间，11位基准副载波相位必须加上180°，从而正弦ROM和余弦 ROM的输出拥有正确的副载波相位（分别为0°和90°）。

对于YIQ处理，在有效视频时间期间，11位基准副载波相位必须加上213°，从而正弦ROM和余弦ROM的输出拥有正确的副载波相位（分别为33°和123°）。

对于所有的方程来说，它们都满足如下条件：

$\omega = 2\pi F_{SC}$

$F_{SC} = 3.579545\text{MHz}(\pm 10\text{Hz})$

YUV颜色空间

正如第8章中所讨论的，色度信号可以由下式表示：

$(U\sin\omega t) + (V\cos\omega t)$

色度幅度为±sqrt($U^2 + V^2$)。

YCbCr颜色空间

如果编码器基于YCbCr颜色空间，那么色度信号可以由下式表示：

$(Cb - 512)(0.504)(\sin\omega t) + (Cr - 512)(0.711)(\cos\omega t)$

至于NTSC-J制式，方程如下：

$(Cb - 512)(0.545)(\sin\omega t) + (Cr - 512)(0.769)(\cos\omega t)$

在这些情况中，正弦ROM和余弦ROM中的值按指示值缩放，以允许调制器复用器直接接收Cb和Cr数据，而不是U和V数据。

YIQ颜色空间

正如第8章中所讨论的，色度信号可以由下式表示：

$(Q\sin(\omega t + 33°)) + (I\cos(\omega t + 33°))$

色度幅度为±sqrt($I^2 + Q^2$)。

❑（B，D，G，H，I，M，N，N_C）PAL

在有效视频期间，CbCr或UV数据调制正弦和余弦副载波，如图9-1所示，结果生成数字色度（C）数据。对于该设计，11位基准副载波相位（参见图9-17）为135°。

对于所有的方程来说，它们都满足如下条件：

$\omega = 2\pi F_{SC}$

对于（B、D、G、H、I、N）PAL：$F_{SC} = 4.433\,618\,75\text{MHz}$（±5Hz）

对于（N_C）PAL：F_{SC} = 3.582 056 25MHz（±5Hz）

对于（M）PAL：F_{SC} = 3.575 611 49MHz（±5Hz）

PAL开关

理论上，因为（sinωt）和（cosωt）副载波是正交的，所以U和V信号能够在解码器中完美地相互分离。但是，如果视频信号中存在失真，例如边带由于低通滤波的非对称衰减，正交性会有所下降，致使U信号和V信号之间产生串扰。

397 ~ 400

除了90°的副载波相位偏移外，PAL还使用V信号的正交行切换来提供U副载波和V副载波之间的频率偏移。当解码时，串扰分量已在正交行载波频率上调制，在纯色区域产生一个称之为Hanover条（Hanover bar）的移动模式。通过梳状滤波器平衡开关行和非开关行之间的等价作用可以抑制该模式。

当PAL开关等于0时，11位基准副载波相位（参见图9-17）和色同步信号相位是相同的（135°）。因此，有效视频期间的11位基准副载波相位必须加上225°，从而正弦ROM和余弦 ROM的输出拥有正确的副载波相位（分别为0°和90°）。

当PAL开关等于1时，11位基准副载波相位加了90°，致使色同步信号的相位为225°。因此，有效视频期间的11位基准副载波相位必须加上135°，从而正弦ROM和余弦ROM的输出拥有正确的副载波相位（分别为0°和90°）。

注意图9-17中，当PAL开关等于1时，生成的是－V副载波，实现的是－V分量。

YUV颜色空间

正如第8章中所讨论的，色度信号由下式表示：

$$(U\sin\omega t) \pm (V\cos\omega t)$$

其中V的符号每行变换一次（称为PAL开关）。

色度幅度为$\pm \mathrm{sqrt}(U^2 + V^2)$。

YCbCr颜色空间

如果编码器基于YCbCr颜色空间，那么（B、D、G、H、I、N_C）PAL的色度信号可以由下式表示：

$$(Cb - 512)(0.533)(\sin\omega t) \pm$$
$$(Cr - 512)(0.752)(\cos\omega t)$$

（M、N）PAL的色度信号可以由下式表示：

$$(Cb - 512)(0.504)(\sin\omega t) \pm$$
$$(Cr - 512)(0.711)(\cos\omega t)$$

在这些情况中，正弦ROM和余弦ROM中的值按指示值缩放，以允许调制器复用器直接接收Cb和Cr数据，而不是U和V数据。

❑ 一般处理

副载波正弦和余弦值的精确度应该最低为9位带符号的正数。调制复用器必须在输出处拥有饱和逻辑，以确保上溢条件和下溢条件分别在最大值和最小值处饱和。

在调制色差信号叠加在一起之后，结果舍入为9位带符号正数。此时，数字调制色度的范围如表9-3所示。生成的数字色度数据结果通过一个拥有相同凸余弦值的消隐信号，以及和亮度数据消隐所用相同的时序信息进行钳制。

401

表9-3　10位数字色度值

视频电平	（M）NTSC	NTSC-J	（B、D、G、H、I、Nc）PAL	（M、N）PAL
峰值色度	328	354	347	328
峰值色同步	112	112	117	117
消隐	0	0	0	0
峰值色同步	－112	－112	－117	－117
峰值色度	－328	－354	－347	－328

❑ 色同步信号生成

如图9-1所示，低通滤波的色差数据和色同步信号包络信息一起复用。在色同步信号时间期间，色差数据应该忽略，而且色同步包络信号插入到Cb、U或者Q通道上（Cr、V或I通道被强制为0）。

色同步包络上升和下降时间应该生成一个凸余弦分布，从而减缓色同步包络的摆率。色同步包络上升和下降时间的典型值为300±100ns。

色同步包络应该足够宽，以生成9个或者10个周期幅度为50%或更大的色同步信息。色同步包络信号乘以正弦ROM的输出，就可以生成色同步信号，它的范围如表9-3所示。

对于专业视频应用来说，色同步信号的相位应该是0°至360°范围内可编程的，从而提供与外部视频信号匹配的可选系统相位。这可通过在色同步时间期间向11位副载波基准相位添加一个可编程值实现（参见图9-17）。

3. 模拟色度（C）生成

数字色度数据可以驱动一个生成0~1.305V输出的10位DAC生成S-video（Y/C）接口的C视频信号。该视频信号在连接器处应该有一个75Ω的源阻抗。

图9-10和图9-11示范了75%彩条的调制色度视频波形。括号中的数字表示满量程输出值为1.305V的10位DAC的数据值。如果DAC不能处理双极视频信号的生成，那么在驱动DAC之前必须在色度数据（以及丢弃的符号信息）上加上一个偏移。在这个实例中，使用的偏移为＋512，将消隐电平定位在10位DAC输出电平的中点上。

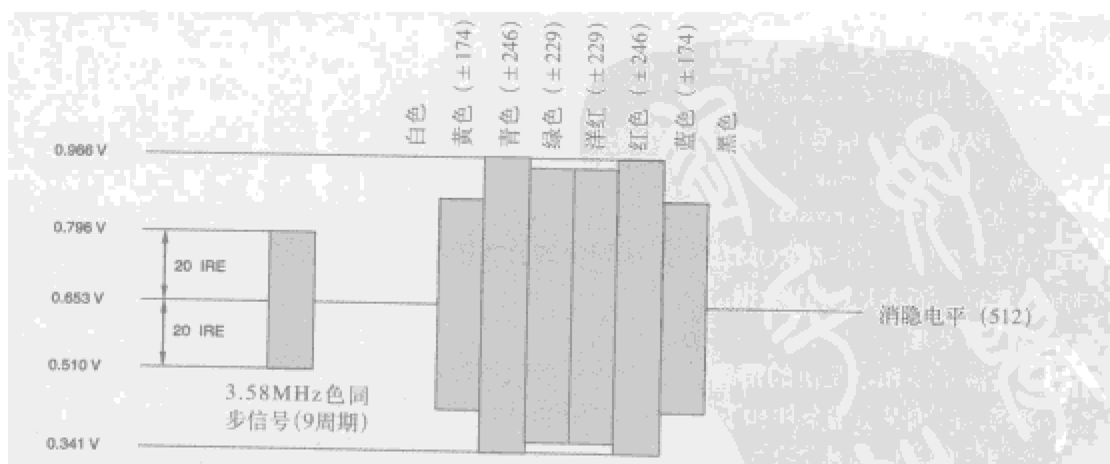


图9-10　75%彩条的（M）NTSC色度（C）视频信号。图中指示的视频电平值为10位

由于DAC的采样/保持动作引入了一个(sinx)/x特性，所以视频数据可以通过一个$[(\sin x)/x]^{-1}$滤波器的数字滤波进行补偿。另外，因为通常在DAC之后提供一个模拟低通滤波器，所以校正可以在模拟滤波器中进行。

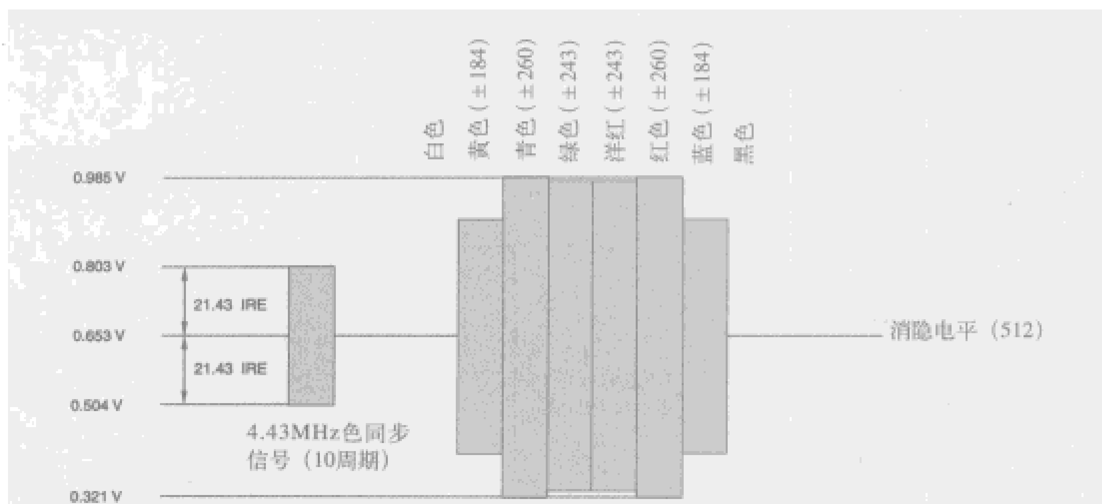


图9-11 75%彩条的（B、D、G、H、I）PAL色度（C）视频信号。图中指示的视频电平值为10位

9.1.5 模拟复合视频

数字亮度（Y）数据和数字色度（C）数据叠加在一起，生成数字复合彩色视频，电平值如表9-4所示。

表9-4 10位数字复合视频电平

视频电平	（M）NTSC	NTSC-J	（B、D、G、H、I、Nc）PAL	（M、N）PAL
峰值色度	973	987	983	973
白色	800	800	800	800
峰值色同步	352	352	369	357
黑色	282	240	252	282
消隐	240	240	252	240
峰值色同步	128	128	135	123
峰值色度	109	53	69	109
同步	16	16	16	16

所生成的结果可以驱动一个生成0~1.305V输出的10位DAC生成复合视频信号。该视频信号在连接器处应该有一个75Ω的源阻抗。

图9-12和图9-13示范了75%彩条的视频波形。括号中的数字表示满量程输出值为1.305V的10位DAC的数据值。

由于DAC的采样/保持动作引入了一个(sin*x*)/*x*特性，所以视频数据可以通过一个[(sin*x*)/*x*]$^{-1}$滤波器的数字滤波进行补偿。另外，因为通常在DAC之后提供一个模拟低通滤波器，所以校正可以在模拟滤波器中进行。

403~404

黑场视频信号

作为一个选项，编码器可以生成一个黑场[或者锁相同步（house sync）]视频信号，用来同步多个视频源。图9-14和图9-15描述了黑场视频信号。注意这些信号和模拟复合视频信号相同，但是不包含任何有效视频信息。括号中的值表示一个满量程输出值为1.305V的10位DAC的数据值。

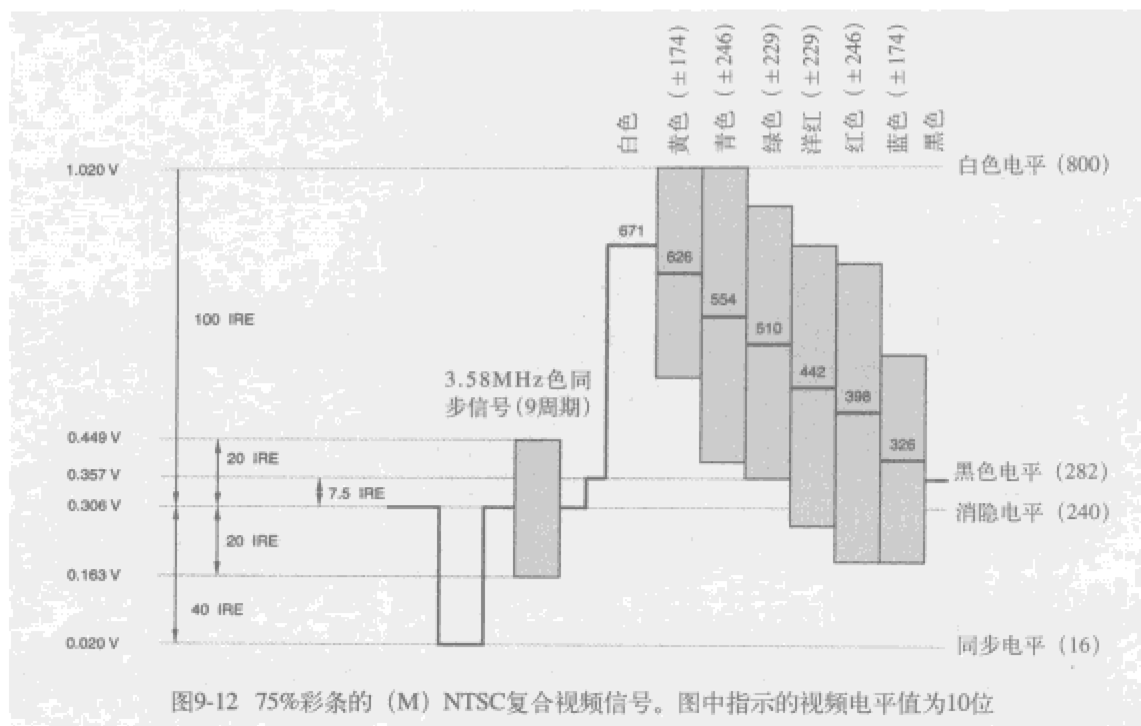
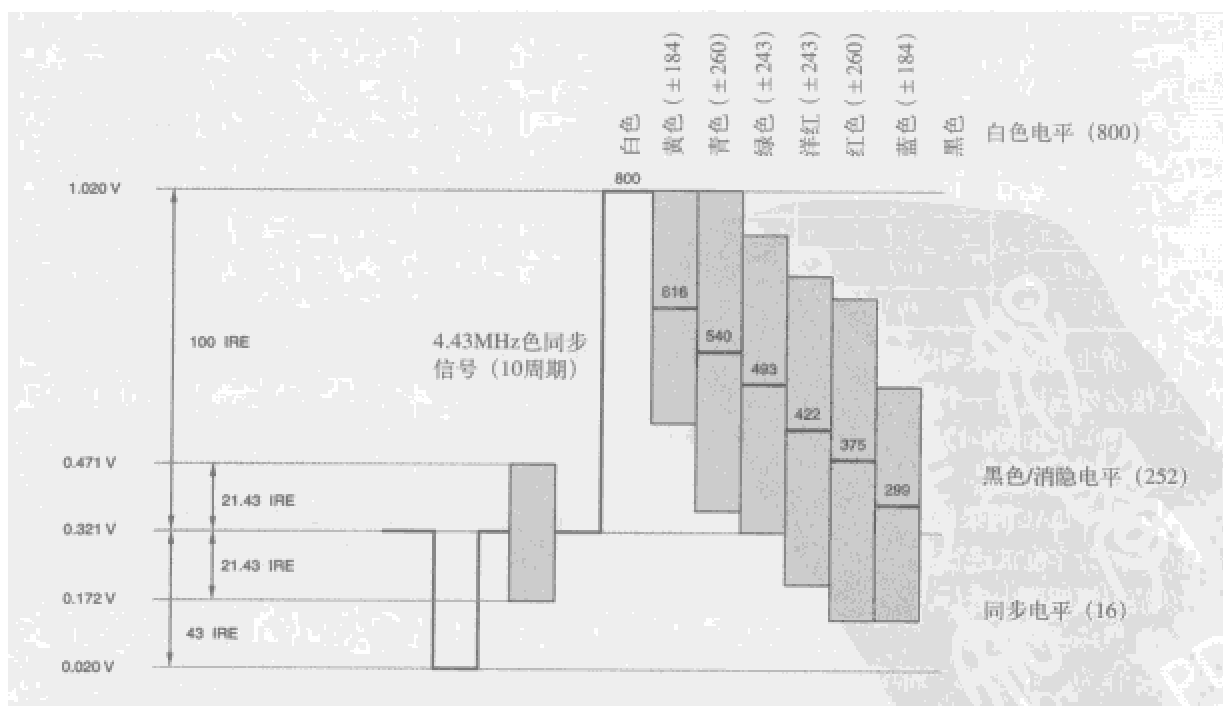


图9-12　75%彩条的（M）NTSC复合视频信号。图中指示的视频电平值为10位



405

图9-13　75%彩条的（B、D、G、H、I）PAL复合视频信号。图中指示的视频电平值为10位

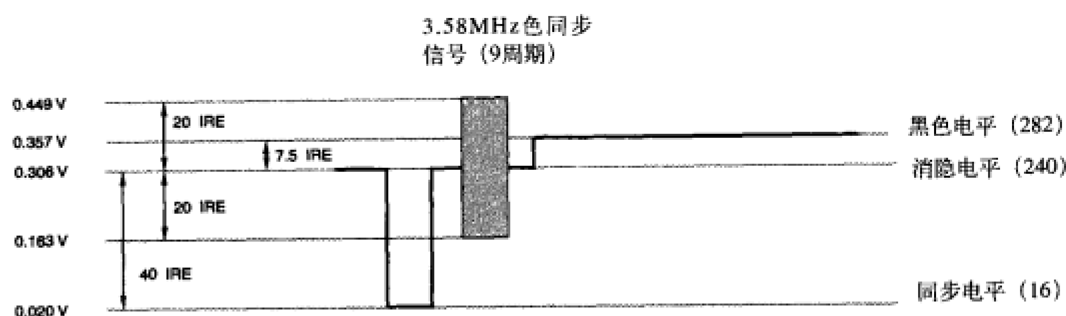


图9-14　（M）NTSC黑场视频信号。图中指示的视频电平值为10位

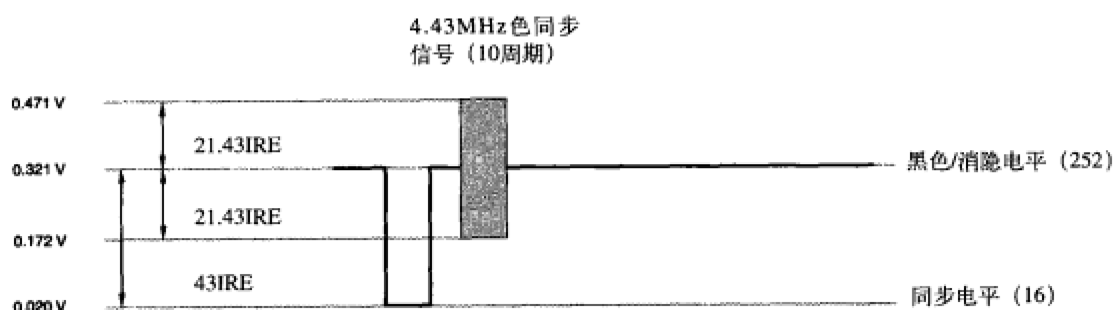


图9-15　（B、D、G、H、I）PAL黑场视频信号。图中指示的视频电平值为10位

406

9.1.6　彩色副载波生成

彩色副载波可以通过使用一个离散时间振荡器（Discrete Time Oscillator，DTO）由采样时钟生成。

当生成的视频用于视频编辑时，维护彩色副载波和同步信息之间的相位关系非常重要。除非相对于同步相位的副载波相位正常维护，否则对视频的编辑可能导致瞬时彩色偏移。PAL还要求一个额外的PAL开关，PAL开关用来每隔一扫描行反转V数据的极性。注意要在编码和解码过程中维护PAL开关的极性。

因为在该设计中，彩色副载波由采样时钟导出，所以采样时钟中的任何抖动都将导致副载波频率的相应抖动。在一些PC中，采样时钟使用一个锁相环（Phase-Lock Loop，PLL）生成，而锁相环没有保持副载波相位抖动低于2°~3°的必要时钟稳定性。

1. 频率关系

❑ （M）NTSC、NTSC-J

正如第8章所示，副载波频率（F_{SC}）和行频（F_H）之间有一个规定的关系：

$F_{SC}/F_H = 910/4$

假定（只是为了示例）采样时钟频率（F_S）为13.5MHz，则有：

$F_S = 858F_H$

将这两个方程结合在一起就得到F_{SC}和F_S之间的关系：

$F_{SC}/F_S = 35/132$

该关系也可以使用采样时钟周期（T_S）和副载波周期（T_{SC}）表示：

$T_S/T_{SC} = 35/132$

彩色副载波相位每个采样时钟必须领先副载波周期这一部分。

- （B、D、G、H、I、N）PAL

正如第8章所示，副载波频率（F_{SC}）和行频（F_H）之间有一个规定的关系：

$F_{SC}/F_H=(1135/4)+(1/625)$

假定（只是为了示例）采样时钟频率（F_S）为13.5MHz，则有：

$F_S=864F_H$

将这两个方程结合在一起就得到F_{SC}和F_S之间的关系：

$F_{SC}/F_S=709\ 379/216\ 000\ 0$

该关系也可以使用采样时钟周期（T_S）和副载波周期（T_{SC}）表示：

$T_S/T_{SC}=709\ 379/216\ 000\ 0$

彩色副载波相位每个采样时钟必须领先副载波周期这一部分。

- （N_C）PAL

在阿根廷使用的（N_C）PAL视频标准中，副载波频率（F_{SC}）和行频（F_H）之间有一个不同

407　的关系：

$F_{SC}/F_H=(917/4)+(1/625)$

假定（只是为了示例）采样时钟频率（F_S）为13.5MHz，则有：

$F_S=864F_H$

将这两个方程结合在一起就得到F_{SC}和F_S之间的关系：

$F_{SC}/F_S=573\ 129/216\ 000\ 0$

该关系也可以使用采样时钟周期（T_S）和副载波周期（T_{SC}）表示：

$T_S/T_{SC}=573\ 129/216\ 000\ 0$

彩色副载波相位每个采样时钟必须领先副载波周期这一部分。

2．正交副载波生成

DTO中包含一个累加器，在累加器中一个较小的数[p]加到另一个数[q]上取模。计数器包含一个加法器和一个寄存器，如图9-16所示。寄存器的内容受到约束，因此如果它们大于等于[q]，那么就将内容减去[q]。加法器的输出信号（X_N）为：

$X_N=（X_{N-1}+p）$模q



图9-16　单级DTO

对于每个时钟周期，加上[p]生成一系列线性递增的数字值。[q]不是[p]的整数倍非常重要，这样才能保证生成的值是连续不同的值，因此上一周期的余数与下一周期的余数不同。

DTO用来将采样时钟频率F_S归并到彩色副载波频率F_{SC}上：

$F_{SC}=(p/q)F_S$

因为[p]拥有有限的字长度，所以DTO输出频率只能以阶（step）变化。当[p]的字长度为[w]时，最小的[p]阶为0.5w，最小的DTO频率台阶为：

$F_{SC}=F_S/2^w$

注意输出频率不能比输入频率的一半大。这意味着输出频率F_{SC}只能按照增量[p]并且在下述范围内变化：

$0<F_{SC}<F_S/2$

在这个应用中，溢出对应着副载波完整周期的结束。

因为只需要余数（表示副载波相位），所以全部完整周期的数量没有用处。在每个时钟周期期间，[q]寄存器的输出展现了在第q个副载波周期中副载波频率的相对相位。通过使用[q]寄存器内容寻址包含一个正弦波特性的ROM，可以生成被采样副载波正弦波的一个数字表示。 408

❑ 单级DTO

使用一个单独的24位或32位模[q]寄存器以及11个最高有效位，可以提供副载波基准相位。这种结构的一个例子如图9-16所示。

❑ 多级DTO

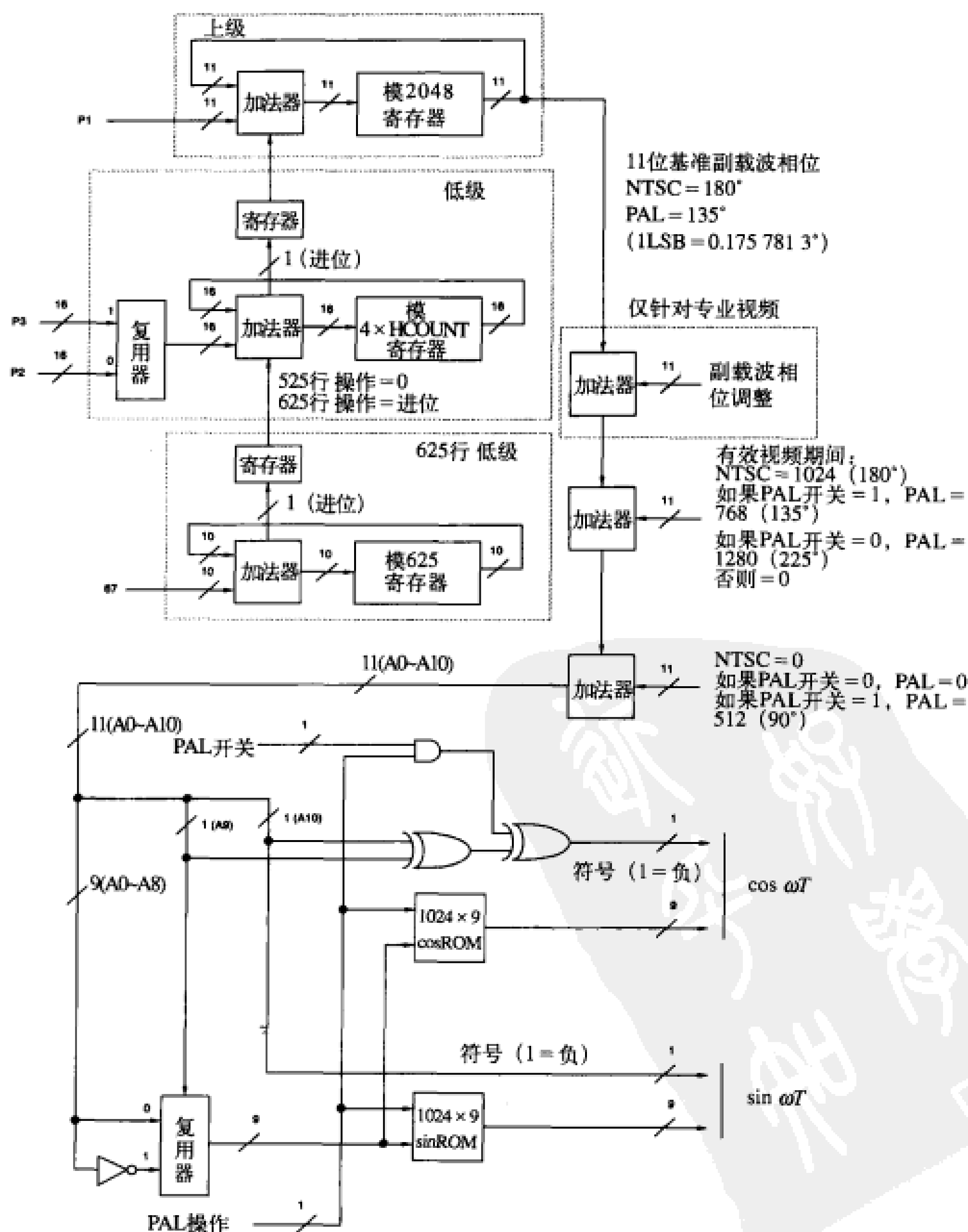


图9-17 3级DTO色度副载波生成

如果将比率划分为2个分数或者3个分数，就可以实现更长的精确度，较高有效分数提供副载波基准相位，如图9-17所示。

为了使用ROM的全部容量，并且使溢出自动化，要求最高有效分数的分母是2的方幂。最低有效分数的分母使用4×HCOUNT来简化硬件计算。

将副载波周期再细分为2048个相位阶，并且使用每个扫描行的总样本数量（HCOUNT），则比率可以如下划分：

$F_{SC}/F_S = (P1 + (P2/(4 \times HCOUNT)))/2048$

P1和P2编程为生成期望的彩色副载波频率（F_{SC}）。模4×HCOUNT计数器和模2048计数器应该在第1场的每个场同步的开头复位，确保生成正确的副载波基准（如图8-5和图8-16所示）。

较低有效级生成一系列的进位，这些进位通过将统计阶变更数值1（即从P1修改为P1+1）纠正上级的近似比率。上级生成一个11位的副载波相位，用来寻址正弦ROM和余弦ROM。

尽管上级加法器自动溢出提供模2048操作，但是下级加法器要求额外的电路，因为4×HCOUNT可能不是（并且通常不是）一个2的方幂整数。在这种情况下，16位寄存器的最大容量为65 535，加法器为所有大于它的值生成一个进位。为了生成正确的进位序列，每次加法器溢出时，有必要调整下一个添加的数，以弥补65 535和4×HCOUNT之间的差。这要求：

$P3 = 65536 - (4)(HCOUNT) + P2$

尽管这修改了下级寄存器的内容，但是进位的序列没有改变，确保了生成正确的相位值。

（M）NTSC操作的P1和P2值使用下述方程确定：

409

$F_{SC}/F_S = (P1 + (P2/(4 \times HCOUNT)))/2048$

$F_{SC}/F_S = (910/4) \times (1/HCOUNT)$

（B、D、G、H、I、N）PAL操作的P1和P2值使用下述方程确定：

$F_{SC}/F_S = (P1 + (P2/(4 \times HCOUNT)))/2048$

$F_{SC}/F_S = ((1135/4) + (1/625)) \times (1/HCOUNT)$

阿根廷使用的（N_C）PAL操作的P1和P2值使用下述方程确定：

$F_{SC}/F_S = (P1 + (P2/(4 \times HCOUNT)))/2048$

$F_{SC}/F_S = ((917/4) + (1/625)) \times (1/HCOUNT)$

因为在P1和P2值计算之后有一个0.1027的余数，所以在625行操作期间使用模625计数器（它的[*p*]值为67）来更加精确地调整副载波的生成。在525行操作期间，进位符号应该总是强制为0。表9-5列举了一些常见的行分辨率、采样时钟频率及相应的HCOUNT、P1和P2值。

表9-5　图9-17中3级DTO的典型HCOUNT、P1和P2值

典型应用	每个扫描行上的总样本（HCOUNT）	4×HCOUNT	P1	P2
13.5MHz（M）NTSC	858	3432	543	104
13.5MHz（B、D、G、H、I）PAL	864	3456	672	2061
12.27MHz（M）NTSC	780	3120	597	1040
14.75MHz（B、D、G、H、I）PAL	944	3776	615	2253

❑ 正弦和余弦生成

无论使用什么类型的DTO，11位副载波相位的每个值对应于2048个波形值中的一个，这些值取自副载波周期期间某个特定点并存储在ROM中。当按相反顺序读取样本值时，采样点应该位于总周期的第4096个点的奇数倍处，以避免边缘效应。

注意ROM中只存储副载波波形的一个象限，如图9-18所示。其他象限的值通过正弦曲线波形

的对称属性生成。使用这一技术的最大相位误差相对于峰－峰幅度为±0.09°（360/2048的一半），在正弦波形信号最陡峭部分，它对应于±0.08%的最大幅度误差。

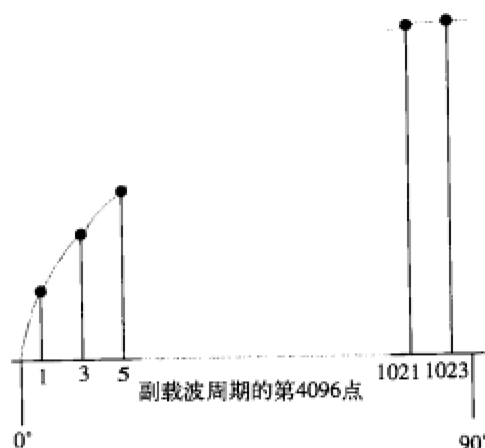


图9-18 副载波周期的一个象限的sinROM和cos ROM中512个存储样本值的位置。其他象限的样本通过反转地址和符号值生成

图9-17还示范了一种由一个11位副载波相位信号生成正交副载波的技术。它使用2个ROM来存储正弦和余弦波形的象限。XOR门反转生成波形时间翻转部分的地址，并且反转输出极性，生成波形的负值部分。另外在符号位中还为V副载波提供一个额外的门，允许插入PAL开关方波，实现V信号在正交扫描行上的相位反转。

9.1.7 行和场时序

行和场计数器用来控制视频时序。

1. 时序控制

为了控制行和场计数器，通常使用独立的行同步（HSYNC#）信号和场同步（VSYNC#）信号。当生成有效视频时，通常使用一个BLANK#控制信号表示。

如果HSYNC#、VSYNC#和BLANK#信号是输入，控制行和场计数器，那么这就称为“从”时序。HSYNC#、VSYNC#和BLANK#信号由系统中的另一个设备生成，由编码器使用生成视频。

410 ~ 411

行和场计数器还可以用于生成基本的视频时序。在这种情况下，这种同步方式称为“主”时序，HSYNC#、VSYNC#和BLANK#从编码器输出，在系统的其他地方使用。

对于BT.656视频接口，使用的是行消隐（H）、场消隐（V）和场（F）信息。在这种应用中，编码器将直接使用H、V和F时序位，而不是再依赖于HSYNC#、VSYNC#和BLANK#控制信号。

表9-6列举了常见采样时钟频率的典型行消隐时序。为了指定何时生成有效视频，使用了一个消隐控制信号（BLANK#）。

表9-6 典型BLANK#信号行时序

典型应用	同步＋后肩消隐（样本）	前肩消隐（样本）
13.5MHz（M）NTSC	122	16
13.5MHz（B、D、G、H、I）PAL	132	12
12.27MHz（M）NTSC	126	14
14.75MHz（B、D、G、H、I）PAL	163	13

2. 行时序

在每个采样时钟的上升边，11位的行计数器都递增1，并由HSYNC#复位。对计数器的值进行监测，以确定何时在每个扫描行上维护及取消各种控制信号，例如色同步包络的起始、色同步包络的结束等。

在从时序操作期间，如果在某行的末尾没有HSYNC#脉冲，那么计数器或者继续递增（推荐），或者自动复位（不推荐）。

3. 场时序

在每个HSYNC#的前沿，10位的场计数器都递增1，并且当VSYNC#和HSYNC#的前沿重合时计数器复位。在场同步中，不是使用一个完全重合的下降沿，而是使用一个大约±64时钟周期的重合窗口（coincident window）以减轻与一些视频时序控制器的交互。如果HSYNC#和VSYNC#的前沿彼此都能在64时钟周期内检测出，那么假定从第1场开始。对计数器的值进行监测，以确定生成哪个扫描行。

412~413

至于隔行（M）NTSC，第1~9扫描行和第264~272扫描行（包含）上的色同步信息应该禁用。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

至于逐行（M）NTSC，第1~9扫描行（包含）上的色同步信息应该禁用。彩色副载波频率上添加一个29.97Hz（30/1.001）的偏移，从而彩色副载波相位相邻场之间是反相的。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

对于隔行（B、D、G、H、I、N、N_C）PAL，在场1、2、5和6期间，扫描行1~6、310~318以及623~625（包括第6行、第318行和第625行）上的色同步信息应该禁用。在场3、4、7和8期间，扫描行1~5、311~319以及622~625（包括第5行、第319行和第625行）上的色同步信息应该禁用。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

对于逐行（B、D、G、H、I、N、N_C）PAL，扫描行1~6和310~312（包括第6行和第312行）上的色同步信息应该禁用。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

对于隔行（M）PAL，在场1、2、5和6期间，扫描行1~8、260~270以及523~525（包括第8行、第270行和第525行）上的色同步信息应该禁用。在场3、4、7和8期间，扫描行1~7、259~269以及522~525（包括第7行、第269行和第525行）上的色同步信息应该禁用。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

对于逐行（M）PAL，扫描行1~8和260~262（包括第8行和第262行）上的色同步信息应该禁用。在剩余的扫描行上，应该启用色同步信息，并在适当的行计数值处禁用色同步信息。

以前的PAL接收机由于不稳定的色同步，在图像的顶部生成彩色的抖动。为了修复这一问题，要实现一种Bruch消隐，以确保每个场同步脉冲后面的第一个色同步的相位是相同的。模拟编码器使用一个曲流门（meander gate）通过偏移场频一行控制色同步的再插入时间。数字编码器简单地跟踪扫描行和场号。现代的接收机不再要求Bruch消隐，但是在确定正在处理哪一个场方面它还是有用的。

在从时序操作期间，如果帧尾没有VSYNC#脉冲，那么计数器或者继续递增1（推荐），或者自动复位（不推荐）。

在主时序操作期间，对于专业视频应用，在625行操作期间可能期望生成2.5扫描行的VSYNC#脉冲。但是，在一些商业视频芯片中，这可能导致第1场与第2场之间的检测问题。

4. 场ID信号

尽管使用HSYNC#和VSYNC#之间的时序关系或者BT.656 F位来指定场1和场2，但是也可以

使用额外的信号指定是生成4场还是8场中的一个场，如表9-7所示。

表9-7 场的编号

FIELD_1信号	FIELD_0信号	HSYNC#和VSYNC#时序关系或BT.656 F位	NTSC场编号		PAL场编号	
0	0	场1	1	奇数场	1	偶数场
0	0	场2	2	偶数场	2	奇数场
0	1	场1	3	奇数场	3	偶数场
0	1	场2	4	偶数场	4	奇数场
1	0	场1	–	–	5	偶数场
1	0	场2	–	–	6	奇数场
1	1	场1	–	–	7	偶数场
1	1	场2	–	–	8	奇数场

在场1、3、5和7期间，应该修改FIELD_0的状态，与VSYNC#的前沿一致。在场1和场5期间，应该修改FIELD_1的状态，与VSYNC#的前沿一致。

对于BT.656视频接口，FIELD_0和FIELD_1可以使用辅助数据发送。 414

9.1.8 干净编码

一般情况下，传统编码器中提供的唯一的滤波器是色差低通滤波器。这会导致亮度分量和色度分量之间存在相当数量的频谱重叠，使解码器端无法完全分离这两个信号。

但是，编码器可以进行额外的处理，减少串色（亮度对色度串扰）和串亮度（色度对亮度串扰）解码器伪影。串色看上去就像一个粗糙的彩条信号图，或者细粒度细节区域中的随机色彩。串亮度看上去就像色度边缘上的一个细粒度图案。

解码器中的串色可以通过在编码器中移除一些高频率的亮度数据降低，高频率亮度数据的移除通过对F_{SC}使用一个陷波滤波器实现。但是，在降低串色的同时，也会丢失亮度细节。

一个更好的办法就是在编码器中对亮度和色度信息使用预梳状滤波器（pre-comb filter）（参见图9-19）。高频率的亮度信息使用预梳状滤波器滤波，使频谱中亮度信息与色度频率的干涉最小。色度信息通过对一些行求平均也进行预梳理滤波，降低串亮度或者悬挂点（hanging dot）模式。

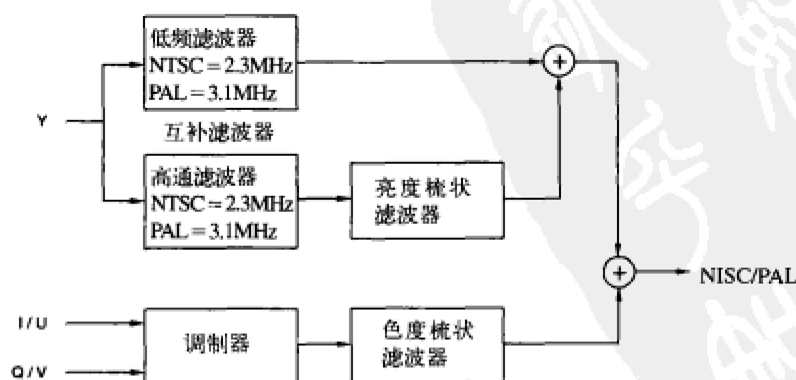


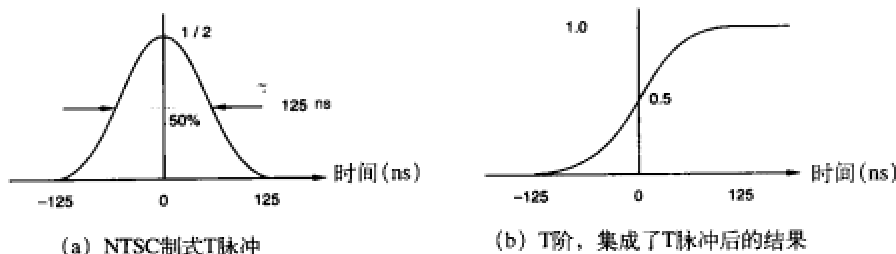
图9-19 干净编码示例

这一技术允许将细粒度的移动亮度（能够在解码器端生成串色）移除，同时保留静态亮度的全部解析度。但是，由于采用了多行平均，所以对角线亮度解析度还是有一小部分的损失。这可

415 以通过改进色度信噪比（Signal-to-Noise Ratio，SNR）进行补偿。

9.1.9 带宽受限边生成

平滑的同步和消隐边可以通过集成一个T脉冲或者凸余弦脉冲生成一个T阶（T step）来生成（图9-20）。NTSC制式使用$T=125$ns的T脉冲；因此，2T阶超过4MHz的信号能量很少。PAL制使用$T=100$ns的T脉冲；在这种实例中，2T阶超过5MHz的信号能量很少。



（a）NTSC制式T脉冲

（b）T阶，集成了T脉冲后的结果

图9-20 带宽受限边生成

T阶可以在一个明确的带宽内提供一个快速的上升时间，且没有振铃。边的10%~90%点之间的上升时间为0.964T。通过为同步边、消隐边和色同步包络选择合适的采样值，这些值可以存储在一个小的ROM中，ROM可以在合适的行计数时触发。通过向前与向后读取ROM的内容，就可416 以生成上升边和下降边。

9.1.10 电平受限处理

某些高度饱和的色彩生成的复合视频电平可能在下游设备中导致问题。

大于100 IRE或者小于－20 IRE（相对于消隐电平）的无效视频电平可以传送，但是可能导致VCR或者解调器失真，并且导致同步分离问题。

大于120 IRE（NTSC）或133 IRE（PAL），或者低于同步脉冲顶部电平的非法视频电平可能无法传送。

尽管在传统的视频应用中这通常并不是一个问题，但是计算机系统通常使用高度饱和的色彩，这有可能生成无效或者非法的视频电平。所以期望可以有选择地将这些信号电平限制在110 IRE附近，在受限的可用色彩和生成合法视频电平之间进行折衷。

进行纠正的一种方法就是调整无效和非法像素的亮度或饱和度，直至达到期望的峰值限制。还有一种方法，即扫描帧缓冲器的内容，标记生成无效或非法视频电平的像素（使用一个单独的重叠面或者色彩变化）。然后用户可以将这个颜色修改为一个更适合的颜色。

在专业的编辑应用中，如果选择在设备之间传送所有的视频信息（包括无效和非法的电平），要求将编辑和处理的伪影减到最小。

9.1.11 编码器视频参数

目前定义了许多工业标准的视频参数来指定NTSC和PAL编码器的相对质量。为了测量这些参数，可以使用视频测试设备监测编码器的输出（生成各种不同的视频测试信号，例如第8章中描述的测试信号）。表9-8示范了家用和演播室质量的编码器的一些典型AC参数值以及一些参数的描述。

表9-8 （M）NTSC和（B、D、G、H、I）PAL编码器的典型AC视频参数

参数	家用级质量		演播室质量		单位
	NTSC	PAL	NTSC	PAL	
微分相位	4		≤1		度
微分增益	4		≤1		%
亮度非线性特性	2		≤1		%
色调准确度	3		≤1		度
色饱和度准确度	3		≤1		%
残留副载波	0.5		0.1		IRE
SNR（遵循EIA-250-C）	48		>60		dB
SCH相位	0±40	0±20	0±2		度
模拟Y/C输出相位偏移	5		≤2		ns
H倾斜	<1		<1		%
V倾斜	<1		<1		%
副载波容限	10	5	10	5	Hz

有几个AC参数，例如组延迟和K因子，依赖于输出滤波器的质量，因此这里不进行讨论。除了本节讨论的AC参数外，还有一些其他参数应该包含在编码器规范中，例如色同步信号频率和容限、行频率、行消隐时间、同步上升和下降时间、色同步包络上升和下降时间、视频消隐上升和下降时间以及YIQ或YUV分量的带宽。

另外还应该指定几种DC参数（例如白色电平和容限、消隐电平和容限、同步高度和容限、峰－峰色同步信号幅度和容限），如表9-9所示。

表9-9 （M）NTSC和（B、D、G、H、I）PAL编码器的典型DC视频参数

参数	家用级质量		演播室质量		单位
	NTSC	PAL	NTSC	PAL	
白色电平（相对于消隐）	714±70	700±70	714±7	700±7	mV
黑色电平（相对于消隐）	54±5	0	54±0.5	0	mV
同步电平（相对于消隐）	－286±30	－300±30	－286±3	－300±3	mV
色同步信号幅度	286±30	300±30	286±3	300±3	mV

1. 微分相位

微分相位失真，通常称为微分相位，指的是色度相位受亮度电平的影响有多大，换句话说，就是当亮度电平变化时色调偏移有多大。相位误差可能是正的也有可能是负的，因此微分相位表示为峰－峰测量值，以副载波相位的度数表示。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级编码器的微分相位参数接近于0.2°或者更小。

417

2. 微分增益

微分增益失真，通常称为微分增益，指的是色度增益受亮度电平的影响有多大，换句话说，就是当亮度电平变化时色饱和度偏移有多大。信号有可能衰减也有可能放大，因此微分增益表示为任意两个电平之间的最大幅度变化，表示为最大色度幅度的百分比。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级编码器的微分增益参数接近于0.2%或者

更小。

3. 亮度非线性特性

亮度非线性特性也称为微分亮度和亮度非线性失真，表示亮度增益受亮度电平的影响有多大，换句话说，就是生成的亮度电平与理想亮度电平之间的非线性关系。

通过使用一个非调制的5级或者10级阶梯测试信号，将最大级和最小级之间的差表示为最大级的一个百分比，用来表示亮度非线性特性。尽管这个参数包含在微分增益和微分相位参数中，但是传统上这个参数单独指定。

4. 色度非线性相位失真

色度非线性相位失真指的是色度相位（色调）受色度幅度（色饱和度）的影响有多大，换句话说，就是当饱和度变化时色调偏移有多大。

通过使用一个调制基底测试信号或者组合测试信号的调制基底部分，可以测量每个色度包和色同步之间的相位差。最大测量值和最小测量值之间的差就是峰－峰值，以副载波相位的度数表示。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

5. 色度非线性增益失真

色度非线性增益失真指的是色度增益受色度幅度（色饱和度）的影响有多大，换句话说，就是指已生成的色度幅度电平和理想色度幅度电平之间的非线性关系，通常看作高饱和度色度信号的衰减。

通过使用一个调制基底测试信号或者组合测试信号的调制基底部分，对测试设备进行调整，从而使中间的色度包为40 IRE。另2个色度包的幅度的测量值和标称值之间的最大差表示色度非线性增益失真，用IRE表示，或者表示为最坏包标称幅度的百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

418～419

6. 色度－亮度互调

色度－亮度互调通常称为交叉调制（cross-modulation），表示亮度电平受色度的影响有多大。这可能是高饱和度色度电平削波或正交失真的结果，而且由于色饱和度的改变还可以表现为一个不规则的亮度变化。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，用理想的50 IRE基底电平和测量亮度电平（在移除色度信息之后）之间的最大差来表示色度－亮度互调，色度－亮度互调可以用IRE表示，或者表示为一个百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

7. 色调准确度

色调准确度表示生成的色调值与理想的色调值之间有多紧密。相位误差可能是正的也有可能是负的，所以色调准确度是最坏情况下的正测量值和最坏情况下的负测量值（相对于标称值）之间的差，以副载波相位的度数表示。这个参数使用EIA或EBU的75%彩条作为测试信号进行测量。

8. 色饱和度准确度

色饱和度准确度表示生成的饱和度值与理想的饱和度值之间有多紧密，该参数使用EIA或EBU的75%彩条作为测试信号。该测量值可能是增益也有可能是衰减，所以色饱和度准确度是最坏情况下的增益测量值和最坏情况下的衰减测量值（相对于标称值）之间的差，表示为标称值的百分比。

9. 残留副载波

残留副载波参数表示白色或灰色信息中副载波信息的量（注意，理想情况下应该没有）。在图像的白色或灰色部分，过多的残留副载波就会引起可见的噪声干扰。

通过使用一个非调制5级或10级阶梯测试信号，有效视频期间副载波的最大峰－峰测量值（用IRE表示）用来指定相对于色同步幅度的残留副载波。

10. SCH相位

SCH（SubCarrier to Horizontal，副载波对行）相位指的是行同步的前沿（位于50%幅度点处）和色同步的0交叉点处之间的相位关系（通过将色同步外插到同步的前沿上）。该误差称为SCH相位，用副载波相位的度数表示。

对于PAL制式，由于同步和副载波频率之间更加复杂的关系，SCH相位的定义稍微有所不同——某一给定行的SCH相位关系每8场只重复一次。因此PAL制式的SCH相位根据EBU Technical Statement D 23-1984（E）定义为"外插到第1场第1行同步脉冲前沿的半幅度点上的色同步的＋U分量的相位"。

当合并2个或者更多个视频信号时，SCH相位非常重要。为了避免彩色偏移或者图像跳动，视频信号必须拥有相同的行、场及副载波时序信息，而且相位必须非常匹配。为了实现这些时序约束，视频信号必须拥有相同的SCH相位关系，因为行同步和副载波是拥有一个明确关系的连续信号。对于编码器来说，允许调整SCH相位来简化2个或更多个视频信号的合并非常普遍。因为NTSC和PAL解码器可能要监测SCH相位来确定要解码哪个彩色场，所以维护正确的SCH相位也非常重要。 420

11. 模拟Y/C视频输出相位偏移

模拟亮度（Y）和色度（C）视频信号之间的输出偏移应该最小，从而避免亮度和色度信息之间的相位偏移误差。当在监视器上观看时，过多的输出相位偏移可以表现为锐利垂直边上可见的伪影。

12. H倾斜

H倾斜也称行倾斜和行时间失真，导致行频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像左边和右边之间亮度的变化。对于数字编码器，例如本章中描述的编码器，H倾斜主要是模拟输出滤波器和传输媒介的伪影。

H倾斜使用一个线条（例如NTC-7 NTSC复合测试信号中的线条）来测量，并且测量倾斜的峰－峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的第一个和最后一个微秒。

13. V倾斜

V倾斜也称为场倾斜和场时间失真，导致场频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像的顶边和底边之间亮度的变化。对于数字编码器，例如本章中描述的编码器，V倾斜主要是模拟输出滤波器和传输媒介的伪影。

V倾斜使用18μs、100 IRE的白条（位于场中心的130行扫描行的中心）或者场方波测量。所测量的是倾斜的峰－峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的前3行和最后3行。

9.1.12　同步锁相支持

在许多实例中，期望能够将编码器的输出同步锁相（调整时序信号）到另一个复合模拟视频信号上，为下游视频处理提供方便。这要求将行、场与彩色副载波频率和相位锁定在一起，详见本章NTSC/PAL解码器一节中的讨论。另外，亮度幅度和色度幅度必须匹配。在同步锁相（genlock）中的一个主要问题就是重新生成的采样时钟可能有过多的抖动，导致彩色伪影。

同步锁相的一个变种就是向编码器发送一个超前的锁相同步［也称为黑场或者超前同步（advance sync）］。该超前补偿从锁相同步生成器到下游处理器（例如混和器）使用的编码器输

出之间的延迟。每个视频源都有它自己的超前锁相同步信号，因此每个视频源都在混合或者处理点进行时间对准。

另一个同步锁相选项允许调整副载波相位，从而副载波可以在混合或处理点与其他视频源匹配。副载波相位必须能够由0°调整到360°。为了在大约10ns台阶中独自定位同步和亮度信息，需要或者总是维护一个0 SCH相位，或者允许进行另一个调整。

421

产品之间的输出延迟变化应该位于大约±0.8ns之内，以允许来自不同同步锁相设备的视频信号正确混合。复用器通常假定2个视频信号完美同步锁相，而且2个视频信号之间过量的时间偏斜将导致不良的混合性能。

9.1.13 alpha通道支持

为专业视频编辑应用设计的编码器可能支持alpha通道。输入的8位或者10位数字alpha数据，以流水方式匹配编码过程的流水线，然后转换为模拟alpha信号（详见第7章讨论）。Alpha数据通常是线性的，所生成的模拟alpha信号（也称为键）的范围为0~100 IRE。alpha通道中没有消隐基底和同步信息。

在支持32位像素的计算机系统中，通常有8位是用于alpha信息的。

9.2 NTSC和PAL数字解码

尽管在NTSC/PAL编码器中，亮度分量和色度分量的组合通常是将两个信号简单地叠加在一起，但是在解码器分离它们时却异常困难。模拟NTSC和PAL解码器已经面世一段时间了。但是，它们的使用还非常困难，在使用过程中要求进行调整，而且提供的视频质量有限。

通过使用数字技术来实现NTSC和PAL解码拥有许多优点，例如方便使用、最少量的模拟调整以及出色的视频质量。数字电路的使用还允许设计更加健壮、更加复杂的Y/C分离器和同步锁相实现。

NTSC/PAL数字解码器的一般框图如图9-21所示。

9.2.1 数字化模拟视频

复合视频信号数字编码的第一步就是使用A/D转换器（A/D Converter，ADC）数字化整个复合视频信号。对于我们的例子来说，使用的是10位的ADC，因此，指示值都是10位的值。

复合和S-Video信号在图9-2、图9-3、图9-10、图9-11、图9-12和图9-13中描述。

视频输入通常是交流耦合的，而且有一个75Ω的AC和DC输入阻抗。因此，在行同步期间，每扫描行视频信号必须被DC复位，以在一个已知的电压电平上定位同步脉冲顶部（sync tip）。

视频信号还必须低通滤波（通常约6MHz），以移除可能导致混叠的任何高频分量。尽管广播的视频带宽受到严格限制，但是家用设备没有这方面的标准。只要愿意，视频源可以生成任意的带宽，接收设备可以接收它所能处理的足够多的带宽。

幅度为0.25倍至2倍的视频信号在家用市场最常见。有效视频和同步信号可以改变幅度，特别是在编辑场合，在这种情况下，视频信号可能由几个不同的视频源合并在一起组合而成。

另外，解码器应该能够处理100%的彩色。但是在实际中可能只有75%的彩色广播，在基带视频中没有这样的限制。对于经常使用的计算机生成的文本和图形，更为常见的是高饱和度彩色。

422

1. DC复位

为了移除视频信号中可能存在的任何DC偏移，并且在一个已知的电平定位它，需要进行DC复位（也称为钳位）。

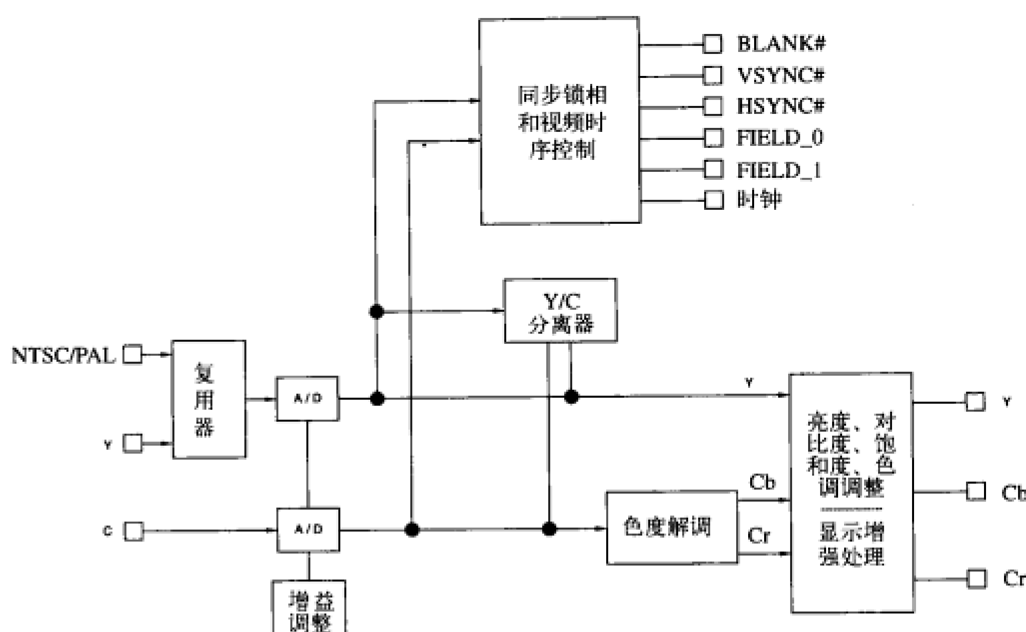


图9-21 典型NTSC/PAL数字解码器实现

423

对于复合或者亮度（Y）视频信号，模拟视频信号被DC复位到每个行同步时间期间的ADC的REF-电压。因此，ADC在同步电平期间生成一个代码0。

对色度（C）视频信号，模拟视频信号被DC复位到行同步时间期间的ADC中间点。因此，ADC在消隐电平期间生成一个代码512。

2. 自动增益控制

使用自动增益控制（automatic gain control，AGC）的目的是为了确保ADC为消隐电平生成一个常量值。如果消隐电平低或者高，那么就将视频信号放大或者衰减，直至消隐电平正确。

在S-Video应用中，应用到亮度视频信号上的增益量要和应用到色度视频信号上的增益量相同。

在DC复位和AGC处理之后，数字化复合和亮度信号加上一个偏移16，以匹配编码器使用的电平。

表9-2、表9-3和表9-4示范了DC复位和自动增益控制完成之后复合和S-Video源的理想ADC值。

❑ 消隐电平确定

确定消隐电平最常见的方法就是对视频信号进行大约0.5MHz的数字低通滤波，移除副载波信息和噪声。然后对后肩进行多次采样，确定平均消隐电平值。

为了限制行－行变化和钳位拖尾（量化误差的结果），结果应该是3~32个连续扫描行的平均结果。另外，后肩电平可以在场消隐期间确定，并且在整个场中使用该结果。

❑ 视频增益选项

在消隐电平与理想消隐电平之间的差处理之后，以下述方式之一生成正确的消隐电平：

(1) 控制一个电压控制的放大器；

(2) 调整ADC的REF＋电压；

(3) 乘以ADC的输出。

在（1）和（2）中，控制增益的模拟信号可以通过DAC或者电荷泵（charge pump）生成。如

果使用的是DAC，那么它的解析度应该是ADC解析度的2倍，以避免量化噪声。出于该原因，电荷泵实现可能更加合适。

选项（2）依赖于ADC能够在很宽的基准电压范围内运转，因此极少实现。

选项（3）极少使用，这是因为数字领域的处理所引入的量化误差。

❑ 同步幅度AGC

这是AGC最常见的模式，当视频信号的特性未知时可以使用它。测量的消隐电平值与理想消隐电平值之间的差用来确定整个视频信号增益的增加量或者减小量。

❑ 色同步幅度AGC

AGC的另一个方法基于色同步幅度。当同步幅度无法与有效视频幅度相关时，在专业视频应用中通常使用这种方法。

首先，不考虑同步脉冲顶部的位置，将消隐电平调整到理想值。这可以通过将视频信号添加或者减去一个DC偏移实现。

接下来，确定色同步幅度。为了限制行之间的变化，色同步幅度的结果可以是3~32个连续扫描行的平均值。

测量的色同步幅度和理想色同步幅度之间的差用来确定整个视频信号增益的增加量或者减小量。在增益调整期间，消隐值不应该改变。

424

❑ AGC选项

对于一些专业视频应用，例如视频信号电平已知是正确的，如果除了同步高度外，所有的视频电平都是正确的，或者如果视频信号中有过多的噪声，那么可能期望禁用自动增益控制。

用户可以指定要使用的AGC值，或者一旦确定之后就将AGC值冻结。

9.2.2　Y/C分离

当解码复合视频时，亮度（Y）和色度（C）必须分离。本章稍后会详细讨论许多完成这一任务的技术。

在Y/C分离之后，Y拥有表9-2中所示的标称值。注意亮度仍然包含同步和消隐信息。调制色度拥有表9-3中所示的标称值。

Y/C分离的质量在由解码器生成的总体视频质量中是一个主要因素。

9.2.3　色差处理

1. 色度（C）解调

色度解调器（图9-22）从Y/C分离器或者色度ADC接收调制色度数据。它生成CbCr、UV或IQ色差数据。

❑（M）NTSC、NTSC-J

在有效视频期间，色度数据使用正弦和余弦副载波数据解调，如图9-22所示，结果生成CbCr、UV或IQ数据。出于该原因，11位基准副载波相位（参见图9-32）和色同步信号的相位是相同的（180°）。

对于YUV和YCbCr处理，在有效视频时间期间，11位基准副载波相位必须加上180°，从而sinROM和cos ROM的输出拥有正确的副载波相位（分别为0°和180°）。

对于YIQ处理，在有效视频时间期间，11位基准副载波相位必须加上213°，从而sinROM和cos ROM的输出拥有正确的副载波相位（分别为33°和123°）。

对于所有的方程来说，它们都满足如下条件：

$\omega = 2\pi F_{SC}$

$F_{SC} = 3.579\ 545\text{MHz}$

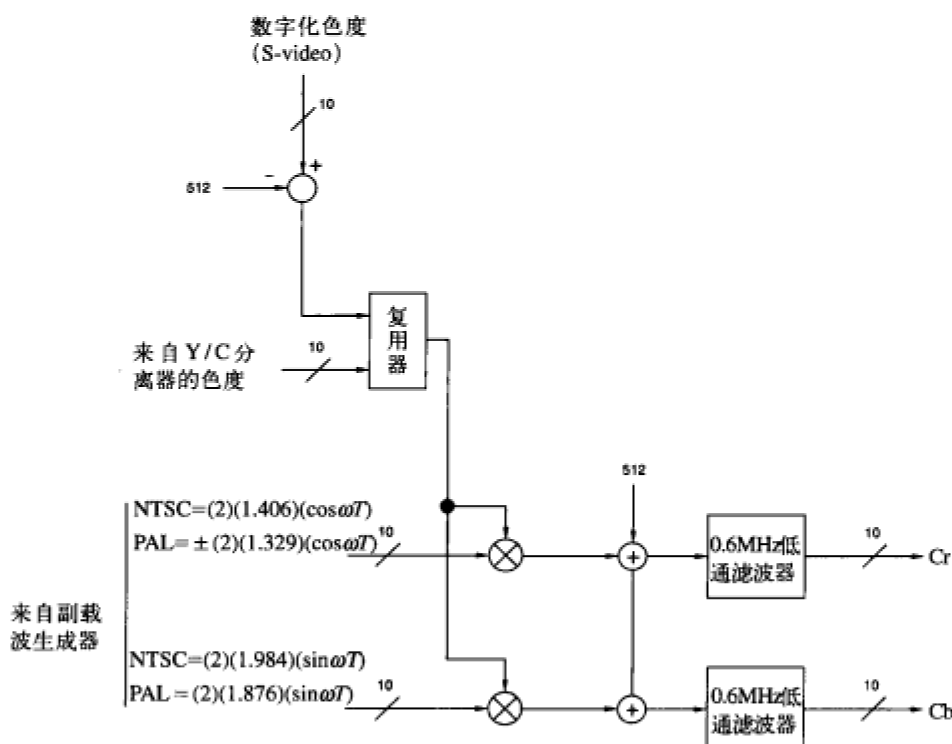


图9-22 直接生成CbCr的色度解调示例

YUV颜色空间处理

正如第8章所示，解调器对色度信号的处理可以表示为： [425]

$(U\sin\omega t) + (V\cos\omega t)$

U通过将色度数据乘以（$2\sin\omega t$）获得，V通过乘以（$2\cos\omega t$）获得：

$((U\sin\omega t) + (V\cos\omega t))(2\sin\omega t)$

$= U - (U\cos 2\omega t) + (V\sin 2\omega t)$

$((U\sin\omega t) + (V\cos\omega t))(2\cos\omega t)$

$= V + (V\cos 2\omega t) + (U\sin 2\omega t)$

$2\omega t$分量被低通滤波移除，从而恢复U信号和V信号。解调器乘法器应该确保上溢条件和下溢条件分别在最大值和最小值处饱和。然后UV信号取整为9位带符号正数并低通滤波。

至于（M）NTSC，U的标称范围为0~±226，V的标称范围为0~±319。

至于日本使用NTSC-J，U的标称范围为0~±244，V的标称范围为0~±344。 [426]

YIQ颜色空间处理

正如第8章所示，对于较旧的解码器，解调器对色度信号的处理可表示为：

$(Q\sin(\omega t + 33^\circ)) + (I\cos(\omega t + 33^\circ))$

解码器的副载波生成器在有效视频期间提供一个33°相位偏移，抵消方程中33°相位这一项。

Q通过将色度数据乘以（$2\sin\omega t$）获得，I通过乘以（$2\cos\omega t$）获得：

$((Q\sin\omega t) + (I\cos\omega t))(2\sin\omega t)$

$= Q - (Q\cos 2\omega t) + (I\sin 2\omega t)$

$((Q\sin\omega t)+(I\cos\omega t))(2\cos\omega t)$

$=I+(I\cos2\omega t)+(Q\sin2\omega t)$

$2\omega t$分量被低通滤波移除，从而恢复I信号和Q信号。解调器乘法器应该确保上溢条件和下溢条件分别在最大值和最小值处饱和。然后IQ信号取整为9位带符号正数并低通滤波。

至于（M）NTSC，I的标称范围为0~±309，Q的标称范围为0~±271。

至于日本使用NTSC-J，I的标称范围为0~±334，Q的标称范围为0~±293。

YCbCr颜色空间处理

如果解码器基于YCbCr颜色空间，那么色度信号可以表示为：

$(Cb-512)(0.504)(\sin\omega t)+(Cr-512)(0.711)(\cos\omega t)$

对于NTSC-J制式，方程为：

$(Cb-512)(0.545)(\sin\omega t)+(Cr-512)(0.769)(\cos\omega t)$

在这些情况中，sinROM和cos ROM中的值通过指示值的倒数缩放，从而允许解调器直接生成Cb和Cr数据，而不是U和V数据。

❑（B、D、G、H、I、M、N、N_C）PAL

在有效视频期间，数字色度（C）数据使用正弦和余弦副载波数据解调，如图9-22所示，结果生成CbCr或UV数据。出于该原因，11位基准副载波相位（参见图9-32）和色同步相位是相同的（135°）。

对于所有的方程来说，它们都满足如下条件：

$\omega=2\pi F_{SC}$

对于（B、D、G、H、I、N）PAL，$F_{SC}=4.433\,618\,75$MHz

对于（NC）PAL，$F_{SC}=3.582\,056\,25$MHz

对于（M）PAL，$F_{SC}=3.575\,611\,49$MHz

通过在Cr或V通道中使用一个交换副载波波形也可以移除PAL开关调制。因此，当PAL并关为逻辑0（色同步相位＝＋135°）时，使用（$+2\cos\omega t$）；当PAL开关为逻辑1（色同步相位＝225°）时，使用（$-2\cos\omega t$）。

YUV颜色空间

正如第8章所示，色度信号可以表示为：

$(U\sin\omega t)\pm(V\cos\omega t)$

427　U通过将色度数据乘以（$2\sin\omega t$）获得，V通过乘以($\pm2\cos\omega t$)获得：

$((U\sin\omega t)\pm(V\cos\omega t))(2\sin\omega t)$

$=U-(U\cos2\omega t)\pm(V\sin2\omega t)$

$((U\sin\omega t)\pm(V\cos\omega t))(\pm2\cos\omega t)$

$=V\pm(U\sin2\omega t)+(V\cos2\omega t)$

$2\omega t$分量被低通滤波移除，从而恢复U信号和V信号。解调器乘法器应该确保上溢条件和下溢条件分别在最大值和最小值处饱和。然后UV信号取整为9位带符号正数并低通滤波。

至于（B、D、G、H、I、N_C）PAL，U的标称范围为0~±239，V的标称范围为0~±337。

对于（M、N）PAL，U的标称范围为0~±226，V的标称范围为0~±319。

YCbCr颜色空间

如果解码器基于YCbCr颜色空间，那么（B、D、G、H、I、N_C）PAL的色度信号可以表示为：

$(Cb-512)(0.533)\sin\omega t\pm(Cr-512)(0.752)\cos\omega t$

(M、N) PAL的色度信号可以表示为：

$$(Cb-512)(0.504)\sin\omega t \pm (Cr-512)(0.711)\cos\omega t$$

在这些情况中，正弦ROM和余弦ROM中的值通过指示值的倒数缩放，从而允许解调器直接生成Cb和Cr数据，而不是U和V数据。

Hanover条

如果本地生成的副载波相位不正确，就会导致一种称为Hanover条的行间模式，在该模式中邻接行对之间一对是实色调误差，一对是互补色调误差。如图9-23所示，该图理想色为绿色，该显示的两个邻接行有一个色调误差（偏向黄色），而下两行为互补色调误差（偏向青色），等等。

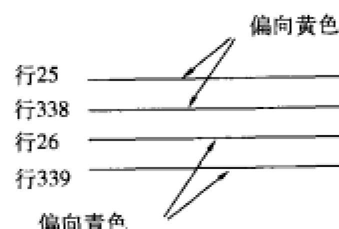


图9-23 Hanover条示例显示。该图中理想色为绿色

这可以通过在本地生成的副载波中引入一个相位误差（θ）来示范：

$$((U\sin\omega t)\pm(V\cos\omega t))(2\sin(\omega t-\theta))$$
$$=(U\cos\theta)-/+(V\sin\theta)$$
$$((U\sin\omega t)\pm(V\cos\omega t))(\pm2\cos(\omega t-\theta))$$
$$=(V\cos\theta)+/-(U\sin\theta)$$

在恒色区域，平均使偶数行和奇数行的作用相等（或者是视觉，或者使用一个延迟行），抵消了正交的串扰分量，只剩下通过（$\cos\theta$）去饱和的真实分量。

2. 低通滤波

解码器要求使用一个比编码器更锐利的滚降滤波器（roll-off filter）来确保对采样混叠分量的足够抑制。注意对于一个13.5MHz的采样频率，3MHz以上的采样混叠分量就非常明显。

图9-24和图9-25示范了（M）NTSC制式解调过程的频谱，该过程与PAL制式相似。在两个图形中，(a) 表示视频信号的频谱，(b) 表示用于解调的副载波的频谱。(a) 和 (b) 的卷积相当于时域中的复用，生成 (c) 中所示的频谱，在这个频谱图中，基带频谱已经移位到F_{SC}和$-F_{SC}$的中间。色度现在成为基带信号，可以通过低通滤波器与位于F_{SC}中间的低频率亮度分离。

解调器后面的低通滤波器是若干个因素之间的折衷。如果只简单地使用一个1.3MHz的滤波器，例如图9-26中所示的滤波器，因为包含了大量的亮度频率，所以会增大串色量。当对大于0.6MHz（NTSC（4.2-3.58））或者1.07MHz（PAL（5.5-4.43））的通带使用低通滤波器时，上边带色度的损失也会引入振铃和色差串扰。如果使用的是一个1.3MHz的低通滤波器，那么它可以包含一些对0.6~1.3MHz之间频率的增益，补偿上边带的部分损失。 428

拥有锐截止（sharp cutoff）的滤波器会加重色度边缘振铃。由于这些原因，通常使用慢速的0.6MHz滚降滤波器，例如图9-27中所示的滤波器。这些滤波器会使色彩解析度较差，但是可以使边上的串色、振铃和色差串扰最小。

如果是在专业视频编辑环境中使用解码器，那么滤波器通带的最大脉动（ripple）应该不超过±0.1dB。为了使滤波器引起的增益和损失伪影累积最小，这是必须的，特别是当同时有多路编码和解码时。

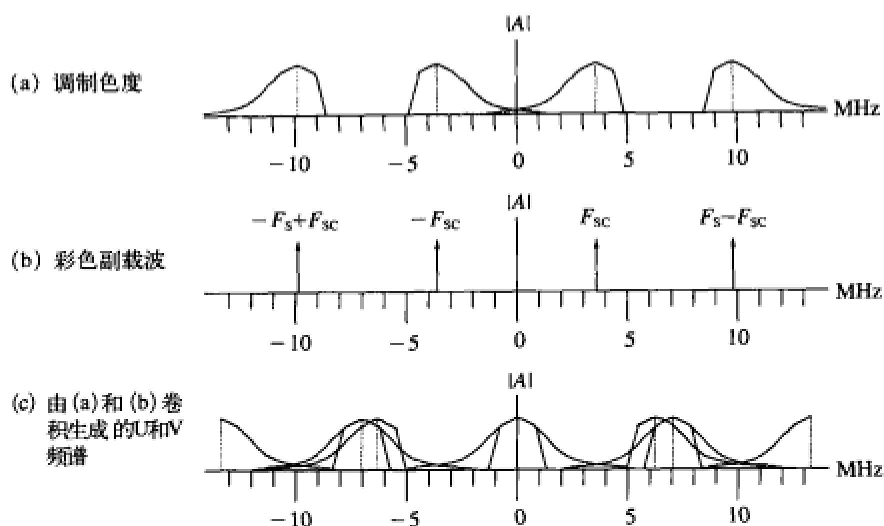


图9-24　NTSC数字色度解调的频谱（F_S＝13.5MHz、F_{SC}＝3.58MHz）

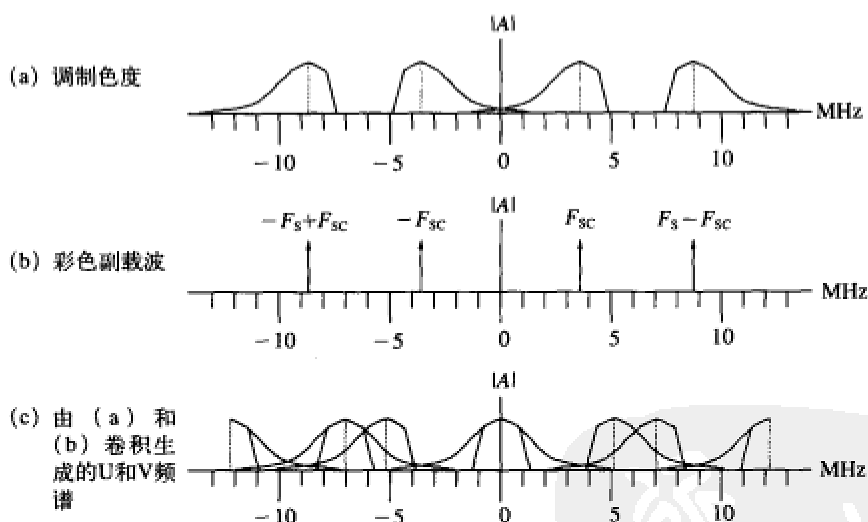


图9-25　NTSC数字色度解调的频谱（F_S＝12.27MHz、F_{SC}＝3.58MHz）

9.2.4　亮度（Y）处理

为了移除同步和消隐信息，Y/C分离器或者亮度ADC的Y数据需要将黑色电平抽取。此时，应该支持负的Y值，从而允许测试信号、抠像信息以及现实世界视频无失真地通过。

中心频率为F_{SC}的陷波滤波器通常是可选的。它可以用来移除Y数据中的任何剩余色度信息。当在PAL制式中使用梳状滤波器滤波Y/C分离时，因为PAL频率包的紧密性，所以陷波滤波器在洁净Y数据方面非常有用。

429

9.2.5　用户调整

1. 对比度、亮度和清晰度

正如第7章中所讨论的，对比度、亮度和清晰度可以实现可编程的调整。另外，色彩瞬变（color transient）改进可以用来提高图像质量。

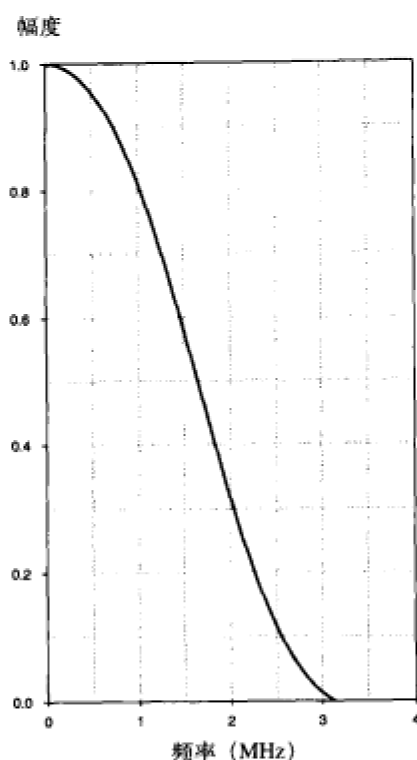


图9-26 典型1.3MHz低通数字滤波器特性

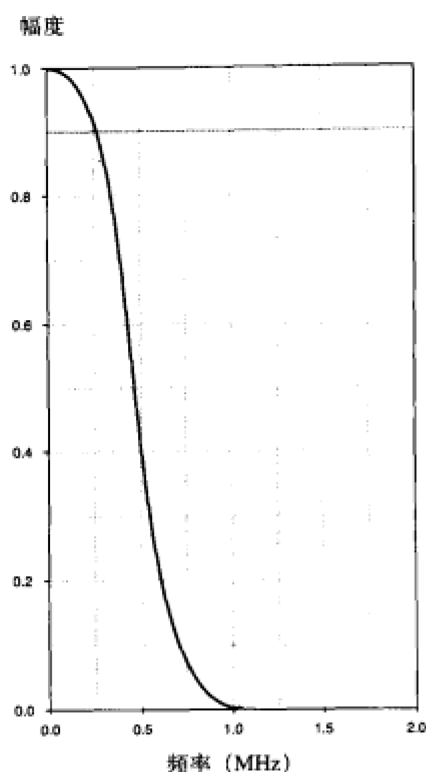


图9-27 典型0.6MHz低通数字滤波器特性

430～431

2. 色调

正如第7章中所讨论的，色调可以实现可编程的调整。

作为选择，为了减少数据路径中的电路，通常将色调调整实现为一个副载波相位偏移，叠加到有效视频时间期间的11位基准副载波相位上（参见图9-32）。结果使正弦和余弦副载波的相位偏移一个常数。11位的色调调整允许从0°到360°以增量0.176°调整色调。

由于PAL解码器中V分量符号的交变，相位偏移（θ）的符号被设置为与V分量的符号相反。相位偏移（θ）的负符号等价于期望相位偏移加上180°。PAL解码器通常不提供色调调整功能。

3. 饱和度

正如第7章中所讨论的，饱和度可以实现可编程的调整。

作为选择，为了减少数据路径中的电路，可以在解调器中实现正弦和余弦值的饱和度调整。

不管是哪一种情况，色同步电平误差信号和用户可编程饱和度值复合在一起，结果用来调整色差信号的增益或者衰减。这里的目的就是为了使色差信号路径中的电路最少。色同步电平误差信号在色同步（以及调制色度信息）没有位于正确的幅度时使用，适当地调整色差信号的饱和度。

有关色同步电平误差信号的更多信息，请参考消色器一节。

4. 自动肤色校正

肤色校正（skin tone correction）可以在NTSC解码器中使用，因为眼睛对肤色非常敏感，而且实际色彩可能在广播过程期间会有些微破坏。如果草地并不是完全正确的绿色，人们可能无法注意到这一点；但是，如果肤色有绿色或者橙色的色彩，人们就无法接受。因为肤色非常接近于+I色轴，所以典型的肤色校正器在一个特定的区域（图9-28）寻找色彩，而且该区域中的任何色

彩都是由接近于肤色的颜色构成的。

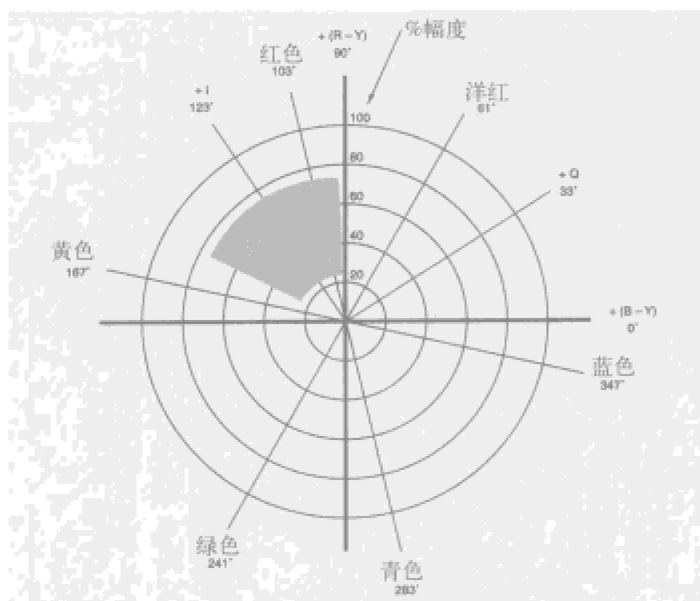


图9-28　典型肤色色彩范围

简单的肤色校正器可以对所有拥有相应+I值的颜色平分Q值。但是，该实现还会改变非肤色颜色。较复杂的实现是，如果颜色的值位于满量程的25%~75%，而且位于+I色轴的±30°之内，那么Q值就可以平分。这样就可以展现肤色区域内接近于理想肤色的任何颜色。

注意各个公司之间肤色的相位角不尽相同。所使用的相位角从116°到126°不等，但是，使用123°（+I色轴）可以简化处理过程。

432

5. 消色器

如果在128个连续扫描行中检测到色同步信号为理想幅度的12.5%或者更小，那么应该将色差信号强制为0。一旦在128个连续扫描行中检测到色同步信号为理想幅度的25%或者更大，那么就应该再次启用色差信号。这种滞后现象防止当色同步幅度位于边界时来回地启用或者禁用色彩信息。

通过强制所有的色同步样本为正，并且多次抽取结果确定一个平均值，可以确定色同步电平。这种平均法所使用的扫描行应该超过3个，以限制行与行之间的变化。

色同步电平误差为理想幅度与平均结果的除值。如果没有检测到色同步，那么应该使用色同步电平误差强制色差信号为0，并且禁用亮度路径上的任何滤波，从而允许输出最大解析度的亮度。

在一些诸如视频编辑的应用中，提供有选择地强制打开或者关闭彩色编码功能这一能力非常有用。

433

9.2.6　颜色空间转换

在由解码器输出之前，YUV和YIQ数据通常要转换为YCbCr或者R′G′B′数据，R′G′B′数据必须在0和1023值处削波，以防止回绕误差。

1. （M）NTSC、（M、N）PAL

❑ YUV颜色空间处理

现在现代的解码器设计都基于YUV颜色空间。对于这些解码器，由YUV生成YCbCr的方程

为：

$$Y_{601}=1.691Y+64$$
$$Cb=[1.984U\cos\theta_B]+[1.984V\sin\theta_B]+512$$
$$Cr=[1.406U\cos\theta_R]+[1.406V\sin\theta_R]+512$$

为了生成位于0~1023范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$$R'=1.975Y+[2.251U\cos\theta_R]+[2.251V\sin\theta_R]$$
$$G'=1.975Y-0.779U-1.146V$$
$$B'=1.975Y+[4.013U\cos\theta_B]+[4.013V\sin\theta_B]$$

为了为专业视频应用生成位于64~940标称范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$$R'=1.691Y+1.928V+64$$
$$G'=1.691Y-0.667U-0.982V+64$$
$$B'=1.691Y+3.436U+64$$

θ_R和θ_B的理想值分别为90°和0°。但是，对于在美国出售的家用电视来说，θ_R和θ_B的理想值通常分别为110°和0°，或者100°和−10°，它们以色彩准确度为代价，以减少微分相位误差的可视性。

❑ YIQ颜色空间处理

对于较老的基于YIQ颜色空间的NTSC解码器设计，由YIQ生成YCbCr的方程为：

$$Y_{601}=1.692Y+64$$
$$Cb=-1.081I+1.664Q+512$$
$$Cr=1.181I+0.765Q+512$$

为了生成位于0~1023范围之内的R'G'B'数据，由YIQ生成R'G'B'的方程为：

$$R'=1.975Y+1.887I+1.224Q$$
$$G'=1.975Y-0.536I-1.278Q$$
$$B'=1.975Y-2.189I+3.367Q$$

为了为专业视频应用生成位于64~940标称范围之内的R'G'B'数据，由YIQ生成R'G'B'的方程为：

$$R'=1.691Y+1.616I+1.048Q+64$$
$$G'=1.691Y-0.459I-1.094Q+64$$
$$B'=1.691Y-1.874I+2.883Q+64$$

❑ YCbCr颜色空间处理

如果设计基于YUV颜色空间，那么通过在解调过程中缩放正弦值和余弦值，或者通过缩放色差低通滤波器系数，可以避免由UV向CbCr的转换。

2. NTSC-J

因为日本使用的（M）NTSC消隐基底为0 IRE，所以颜色空间转换方程与标准的（M）NTSC稍微有所不同。 434

❑ YUV颜色空间处理

现在现代的解码器设计都基于YUV颜色空间。对于这些解码器，由YUV生成YCbCr的方程为：

$$Y_{601}=1.564Y+64$$
$$Cb=1.835U+512$$
$$Cr=[1.301U\cos\theta_R]+[1.301V\sin\theta_R]+512$$

为了生成位于0~1023范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$$R'=1.827Y+[2.082U\cos\theta_R]+[2.082V\sin\theta_R]$$
$$G'=1.827Y-0.721U-1.060V$$

$B' = 1.827Y + 3.712U$

为了为专业视频应用生成位于64~940标称范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$R' = 1.564Y + 1.783Y + 64$

$G' = 1.564Y - 0.617U - 0.908V + 64$

$B' = 1.564Y + 3.179U + 64$

θ_R的理想值为90°。但是，对于在日本出售的电视机来说，θ_R的理想值通常为95°，以色彩准确度为代价，减少微分相位误差的可视性。

❑ YIQ颜色空间处理

对于较老的基于YIQ颜色空间的NTSC解码器设计，由YIQ生成YCbCr的方程为：

$Y_{601} = 1.565Y + 64$

$Cb = -1.000I + 1.539Q + 512$

$Cr = 1.090I + 0.708Q + 512$

为了生成位于0~1023范围之内的R'G'B'数据，由YIQ生成R'G'B'的方程为：

$R' = 1.827Y + 1.746I + 1.132Q$

$G' = 1.827Y - 0.496I - 1.182Q$

$B' = 1.827Y - 2.024I + 3.115Q$

为了为专业视频应用生成位于64~940标称范围之内的R'G'B'数据，由YIQ生成R'G'B'的方程为：

$R' = 1.564Y + 1.495I + 0.970Q + 64$

$G' = 1.564Y - 0.425I - 1.012Q + 64$

$B' = 1.564Y - 1.734I + 2.667Q + 64$

❑ YCbCr颜色空间处理

如果设计基于YUV颜色空间，那么通过在解调过程中缩放正弦值和余弦值，或者通过缩放色差低通滤波器系数，可以避免由UV向CbCr的转换。

3. (B、D、G、H、I、N_C) PAL

❑ YUV颜色空间处理

由YUV生成YCbCr的方程为：

$Y_{601} = 1.599Y + 64$

$Cb = 1.875U + 512$

$Cr = 1.329V + 512$

为了生成位于0~1023范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$R' = 1.867Y + 2.128V$

$G' = 1.867Y - 0.737U - 1.084V$

435

$B' = 1.867Y + 3.793U$

为了为专业视频应用生成位于64~940标称范围之内的R'G'B'数据，由YUV生成R'G'B'的方程为：

$R' = 1.599Y + 1.822V + 64$

$G' = 1.599Y - 0.631U - 0.928V + 64$

$B' = 1.599Y + 3.248U + 64$

❑ YCbCr颜色空间处理

通过在解调过程中缩放正弦值和余弦值，或者通过缩放色差低通滤波器系数，可以避免由UV向CbCr的转换。

9.2.7 同步锁相

同步锁相电路的目的就是为了从视频信号中恢复采样时钟和时序控制信号（例如行同步、场同步和彩色副载波）。因为原始的采样时钟不可用，所以通常通过使用一个锁相环将水平行频F_H乘以每行期望的采样数量从而生成采样时钟。另外，彩色副载波必须重新生成，并且锁定到被解码视频信号的彩色副载波上。

但是，这里还有几个问题。视频信号可能包含噪声，使同步边的确定不可靠。每行上行同步边之间的时间量可能稍微有所不同，特别是在模拟视频磁盘录像机（videotape recorder，VCR）中，由于机械方式的限制，不同更加明显。对于模拟VCR，瞬时行与行之间的偏差可以达到±100ns，每场开头与末尾之间的行偏差达到±5μs。当模拟VCR位于特殊功能模式中时，例如快进或静止图像，行同步信号之间的时间量与标称值之间的偏差可以达到±20%。

和行同步信息一样，场同步信息也必须恢复。然而，模拟VCR，除了破坏了SCH相位关系外，还在场边界执行头切换（head switching）。场边界通常位于有效视频的末尾和场同步起始之间。当头切换发生时，一个视频信号（场n）被另一个视频信号（场$n+1$）替换，而后一个视频信号与第一个视频信号之间的时间偏差未知。每个场的场时序中的偏差可能达到±1/2行。最终，可能会生成比普通行或场同步长的行或场同步。

通过检测水平行时序，有可能自动地确定视频源是处于普通模式还是处理特殊功能模式。在普通模式操作期间，水平行时间在整场中的典型变化通常不超过±5μs。位于±5μs窗口之外的行时序可以用来启用特殊功能模式时序。在检测算法中应该使用滞后作用，以防止在视频时序位于两种模式的边界上时，在普通模式和特殊功能模式两种操作中不停地来回切换。执行行同步和场同步检测的典型电路如图9-29所示。

在没有视频信号时，解码器应该能够有选择地自由运行，继续为系统生成视频时序信息，不遗失一个节拍。在没有输入信号时，应该禁用任何自动增益电路，而且解码器应该提供选项，即或者选择透明（从而可以检控输入源），或者选择自动冻结输入数据（以补偿短期的信号丢失），或者选择自动停播（autoblack）输出数据（以避免驱动混频器或VCR的潜在问题）。 436

1. 行同步检测

以前的解码器通常使用模拟同步限幅技术，以确定同步脉冲前沿的中点，并且使用PLL将行频率放大为采样时钟频率。但是，由于模拟同步限幅器的准确度不高，再加上PLL的稳定性有限，从而导致采样时钟抖动和噪声放大。当在Y/C分离中使用梳状滤波器时，视频数据写入和读取之间的长延迟意味着即使是一个小的采样时钟频率误差，也会导致延迟占副载波周期的很大一个比例，使梳状滤波器的有效性不起作用。

❑ *粗粒度行同步锁定*

粗粒度行同步锁定能够实现较快速的锁定时间。数字化的视频通过0.5MHz的低通滤波，移除高频信息，例如噪声和彩色副载波信息。在低通滤波的数据上执行同步检测还可以在提供快速同步边（上升和下降时间小于一个时钟周期）的情况下提供边的定形。

每个采样时钟周期，11位的行计数器都将递增1，并且在达到HCOUNT值时复位为0x001，其中HCOUNT指定每行采样的总数。值0x001表示期望行同步的开头。当行计数器的值为（HCOUNT－64）时，启用同步门（sync gate），允许检测恢复的同步信息。

在时钟频率校正或者其他调整之前，应该检测到多至5个连续的丢失同步脉冲。一旦检测到同步信息，就禁用同步门，直至下一次行计数器的值为（HCOUNT－64）。这有助于滤波器滤除噪声、齿状和均衡脉冲。如果恢复的行同步的前沿没有位于所期望的±64时钟周期（近似为

±5μs）之内，那么行计数器就要复位为0x001，从而更精密地重新对齐边。

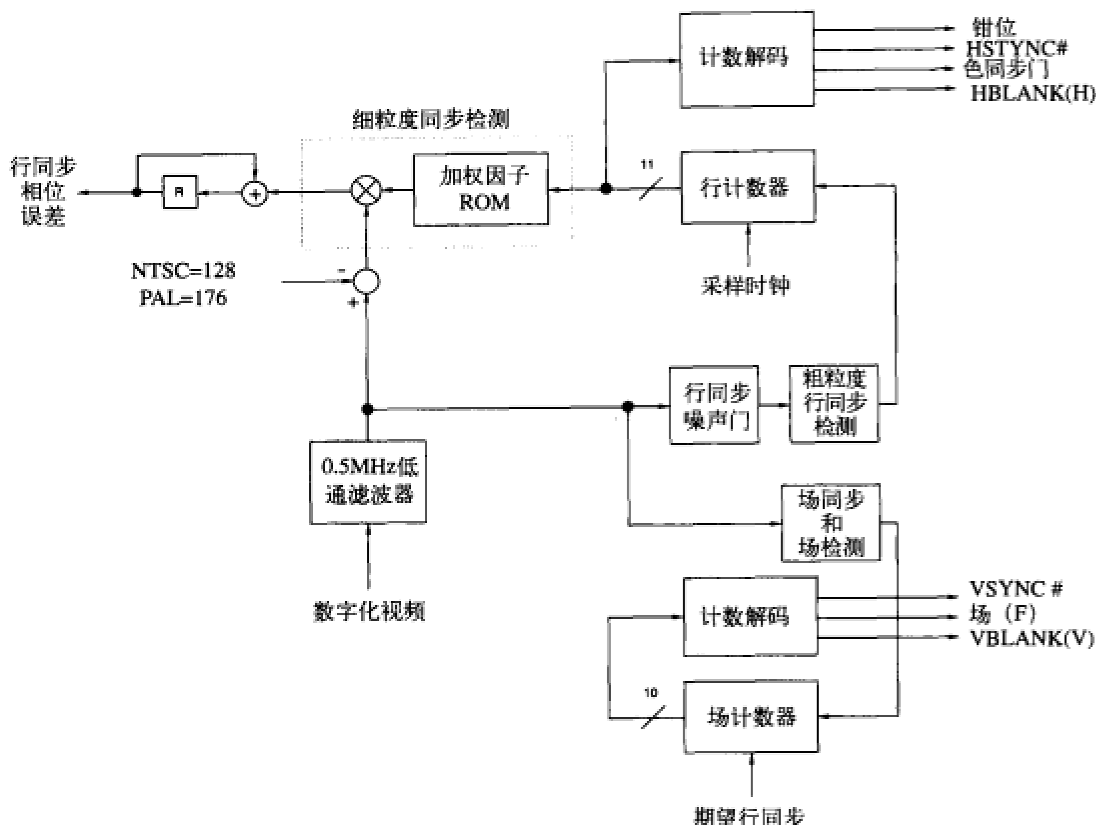


图9-29 同步检测和相位比较器电路

为了监控恢复的行同步脉冲的宽度，还需要包含额外的电路。如果行同步脉冲不是近似地为正确的脉冲宽度，忽略它，将它视为旁路（missing）同步脉冲。

如果恢复的行同步的前沿位于期望的±64时钟周期（近似为±5μs）之内，那么需要使用细粒度的行同步锁定电路细调时序信息。

❑ 细粒度行同步锁定

从0.5MHz低通滤波的视频数据中抽取1/2同步幅度，从而同步时序基准点（50%同步幅度）位于0处。

行同步前沿可以通过将同步边区域内的加权样本加在一起确定。为了执行滤波，加权因子由一个由行计数器触发的计数器从ROM中读取。当中心加权因子（A0）与同步前沿的50%幅度点一致时，结果合计为0。典型的加权因子包括：

437
438

A0＝102/4096

A1＝90/4096

A2＝63/4096

A3＝34/4096

A4＝14/4096

A5＝5/4096

A6＝2/4096

这种安排需要使用同步边的许多时序信息，并且要抑制噪声。注意要包含该电路，从而避免

处理行同步的后沿。

图9-30示范了细粒度同步相位比较器的操作。图9-30a示范了NTSC制式的同步前沿。图9-30b示范了生成的加权因子，当加权因子乘以同步信息时，所生成的波形如图9-30c所示。当A0系数与同步的50%幅度点一致时，波形合计为0。同步边的失真，导致锁定点有轻微偏移，但是可以通过低通滤波最小化，从而在处理之前有效地定形同步边。

❑ 采样时钟生成

图9-29中的行同步相位误差信号用来调整行锁定PLL的频率，如图9-31所示。行锁定PLL通常在每行上生成一个恒定数量的时钟周期，而不考虑任何行时间偏差。PLL的自由运行频率应该是所要求的标称采样时钟频率（例如13.5MHz）。

使用一个基于VCO的PLL拥有这样的优点，即可以在一个较大范围内调整采样时钟频率，在处理标称视频规范之外的视频时序变化方面非常有用。但是，这种方法的缺点是，由于采样时钟的抖动，可能会存在可见的色调伪影，而且Y/C分离质量不佳。

使用一个基于VCXO的PLL拥有这样的优点，即采样时钟抖动最小。但是，采样时钟频率的范围只能在一个小范围内调整，限制解码器处理非标准视频时序的能力。

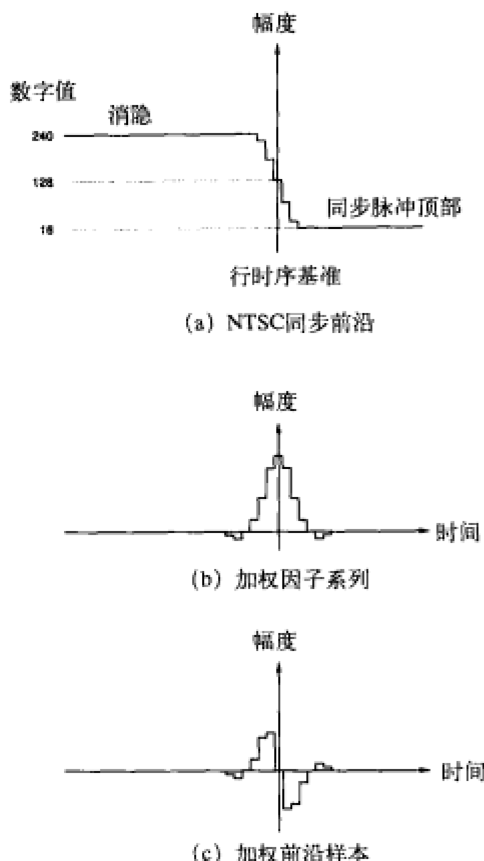


图9-30 细粒度锁定相位比较器波形

理想情况下，无论是哪一种设计，采样时钟的上升边与行同步前沿1/2幅度点对齐，而且通常生成一个每行采样时钟周期数量（HCOUNT）固定的采样时钟。

另一种方法就是使用一个固定频率的时钟（例如13.5MHz）对视频信号进行异步采样。因为在这种情况下，采样时钟没有与行同步对齐，所以在实际采样位置和理想采样位置之间存在一个相位差。对于传统的同步锁相方案，这个相位差由恢复的行同步和期望的行同步之间的差确定。

理想的采样位置定义为与行锁定PLL生成的采样时钟对齐。行同步相位误差信号用来控制数据的两个样本之间的插值，而不是通过控制采样时钟频率，生成理想的采样值。如果在Y/C分离中使用梳状滤波，可以插值数字化的复合视频，生成理想的采样点，通过更精确地对齐样本而提供更好的Y/C分离。



图9-31 典型行锁定采样时钟生成

2. 场同步检测

数字化的视频通过0.5MHz的低通滤波，移除高频信息，例如噪声和彩色副载波信息。每个

439 期望的行同步，10位的场计数器都将递增1，并且在达到525或者625时复位为0x001。值0x001表示期望第1场场同步的开头。

可以对场同步期间的末尾进行检测，根据操作模式设置场计数器的值。通过监测恢复的场同步和行同步之间的关系，可以检测场1和场2之间的对比信息。如果恢复的行同步发生超过64时钟周期，但是小于（HCOUNT/2）时钟周期，那么期望行同步之后的时钟周期、场计数器不进行调整，以避免场计数器两次递增。如果在场计数器递增之后，恢复的行同步发生了（HCOUNT/2）或者更多时钟周期，那么场计数器再次递增。

在特殊功能操作期间，场和行时序信息之间不再有任何关系，所以无法检测场1和场2之间的对比关系。因此，为了合成场1和场2的时序信息，每隔一次场同步末尾的检测，都应该相应地设置场计数器。

3. 副载波生成

至于编码器，彩色副载波通过使用DTO（图9-32）由采样时钟生成，而且也适用编码器一节中所讨论的频率关系。

与编码器不同，所生成的副载波的相位必须持续调整，以匹配被解码的视频信号。

副载波锁定电路相位比较生成的副载波和即将到来的副载波，结果生产一个F_{SC}误差信号，指示相位误差的量。F_{SC}误差信号加到P值上，以连续地调整DTO阶的大小，从而调整所生成的副载波相位，匹配被解码的视频信号。

因为使用一个22位的单阶DTO来分隔采样时钟，生成图9-32所示的副载波，所以P值由下述方程确定：

$$F_{SC}/F_S = (P/4\ 194\ 303) = (P/(2^{22}-1))$$

其中F_{SC}＝期望副载波频率，F_S＝采样时钟频率。流行采样时钟频率的P值如表9-10所示。

表9-10 图9-32中单阶22位DTO的典型HCOUNT值和P值

典型应用	每个扫描行上的总样本（HCOUNT）	P
13.5MHz（M）NTSC	858	1 112 126
13.5MHz（B、D、G、H、I）PAL	864	1 377 477
12.27MHz（M）NTSC	780	1 223 338
14.75MHz（B、D、G、H、I）PAL	944	1 260 742

❑ 副载波锁定

副载波锁定电路（图9-33）的目的是将生成的彩色副载波相位锁定到被解码视频信号的彩色副载波上。

440~441 数字复合视频（或者数字色度视频）将消隐电平抽除。另外数字或色度视频还通过色同步门进行门控，确保值为0的数据位于色同步时间之外。色同步门信号应该定时消除色同步的边，因为色同步的边可能有瞬态失真，会降低相位测量的准确度。

色同步数据与本地生成的色同步进行相位比较。注意信号信息也必须进行比较，这样锁定就不会发生在180°异相信号上。色同步累加器平均16个样本，并且对两个邻接行的累积值求平均以生成误差信号。当本地副载波正确定相时，就取消交变行的累积值，相位误差信号为0。误差信号以行频采样，并由环路滤波器处理，环路滤波器应该设计为实现大约10行（对于噪声视频信号可能需要50行或者更多行）的上锁时间。另外，由于没有色同步，所以还期望在各个场期间更新误差信号。所得到的结果F_{SC}误差信号用来调整生成本地副载波的DTO（图9-32）。

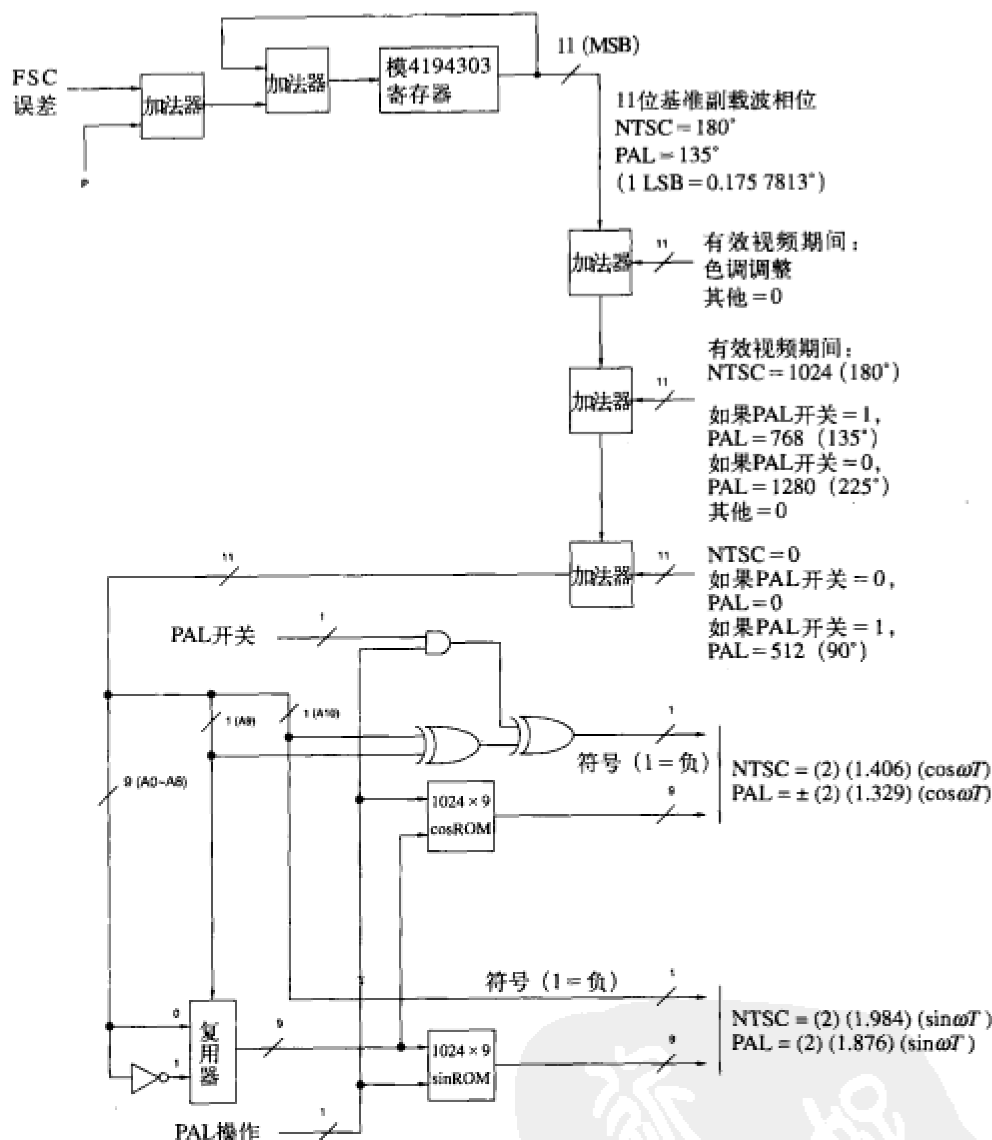


图9-32 色度副载波生成

在PAL操作期间，鉴相器（phase detector）也恢复在生成交换V副载波过程中使用的PAL开关信息。PAL开关D双稳态多谐振荡器（flip-flop）通过比较本地开关传感器和累积色同步值的符号，与传递进来的信号同步。如果传感器有16行始终不正确，那么双稳态多谐振荡器就会复位。

注意副载波锁定电路应该能够处理±200Hz的短期频率偏差（少数几个帧的）、±500Hz的长期频率偏差以及（少数几个帧的）25~200%色同步幅度标称值的多至5%的短期幅度偏差。10行的上锁时间需要能够调节没有正确编辑（即没有正确处理SCH相位关系）的视频信号，或者由于定格（freeze-framing）、特技等所产生的非标准视频信号。这10行能够使副载波在有效视频时间之前锁定，确保在图像的开头色彩表示正确。

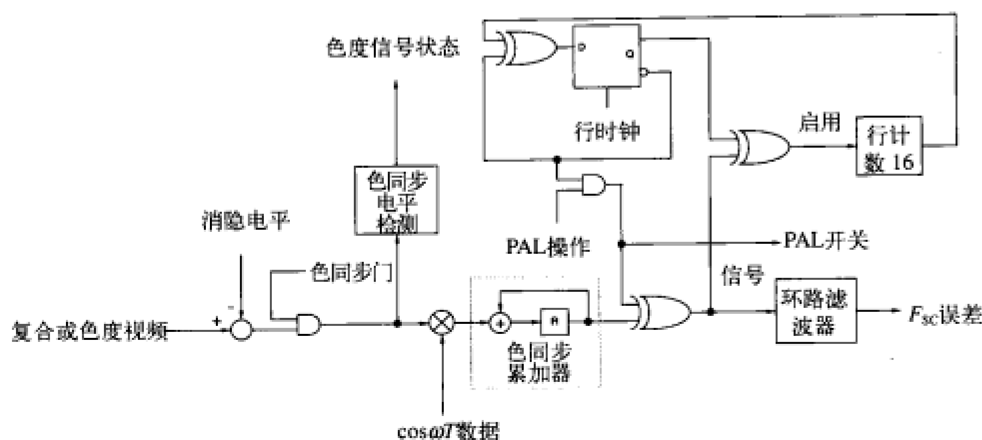


图9-33　副载波相位比较器电路

9.2.8　视频时序生成

1. HSYNC#（行同步）生成

在采样时钟的每个上升边，11位行计数器都递增1。为了确定什么时候生成色同步门、HSYNC#输出、行消隐等信息，需要对这个计数进行监测。一般情况下，每次将这个计数器复位为0x001时，就声称一次HSYNC#输出。HSYNC#的准确时序取决于所使用的视频接口，详见第6章的讨论。

❑ H（行消隐）生成

为了指定行消隐期什么时候发生，可以实现一个行消隐信号H。H的时序取决于所使用的视频接口，详见第6章的讨论。

通过将起始和停止消隐寄存器集成在一起，用户可以对行消隐时序进行编程控制。这些寄存器的值与行计数器的值进行比较，用来声称和取消H控制信号。

2. VSYNC#（场同步）生成

442
~
444

在HSYNC#的每个上升边，10位场计数器都递增1。一般情况下，每次将这个计数器复位为0x001时，就声称一次VSYNC#输出。VSYNC#的准确时序取决于所使用的视频接口，详见第6章的讨论。

❑ F（场）生成

为了指定解码的是场1还是场2，可以实现一个场信号F。F的准确时序取决于所使用的视频接口，详见第6章的讨论。

当解码模拟VCR的输出并且VCR处于特技模式（例如静止和快进）时，已经不再拥有足够的时序信息来确定场1和场2之间的时序。因此，通过VSYNC# / HSYNC#关系（或者F信号）指定的场1和场2的时序应该合成，而且还有可能无法反映被解码视频信号的真实场时序。

❑ V（场消隐）生成

为了指定场消隐期什么时候发生，可以实现一个场消隐信号V。V的准确时序取决于所使用的视频接口，详见第6章的讨论。

通过将起始和停止消隐寄存器集成在一起，用户可以对场消隐时序进行编程控制。这些寄存器的值与场计数器的值进行比较，用来声称和取消V控制信号。

3. BLANK#生成

复合消隐信号BLANK#是H信号和V信号的逻辑NOR。

当声称BLANK#时，可以强制RGB数据为0值。YCbCr数据可以被强制为8位的值，Y数据被强制为16位的值，Cb和Cr数据被强制为128位的值。另外，RGB或者YCbCr数据输出可以不是空的，允许输出场消隐期（VBI）数据，例如隐藏字幕、图文电视、宽屏幕信令以及其他信息。

4. 场标识

尽管行同步（HSYNC#）信号和场同步（VSYNC#）信号之间的时序关系或者F信号可以用来指定解码的是场1还是场2，但是使用1个或者2个额外的信号可以指定解码的是4场或者8场中是哪一场，如表9-7所示。我们称这些额外的控制信号为FIELD_0和FIELD_1。

在场1、3、5和7期间，FIELD_0应该在VSYNC#的开头改变状态，或者与F一致。在场1和场5期间，FIELD_1应该在VSYNC#的开头改变状态，或者与F一致。

❑ NTSC场标识

场1和场3的开头可以通过监测副载波相位与同步之间的关系确定。如图8-5所示，在场1的开头，相对于同步，理想的副载波相位为0°；在场3的开头，相对于同步，理想的副载波相位为180°。

在现实世界中，SCH相位关系中存在容限。例如，尽管理想的SCH相位关系在信源处可能非常完美，但是视频信号通过同轴电缆传输以后，可能会由于电缆特性，导致SCH相位关系产生偏移。因此，所使用的相位应该是理想相位再加上或者减去一个容限。尽管视频标准中指定±40°（NTSC）或者±20°（PAL）是一个可接受的容限，但是许多解码器设计使用多至±80°的容限。

当检测到SCH相位关系没有位于正确的容限范围时，解码器应该继续解码，就像没有发生错误一样。如果这一条件持续了若干帧，表示视频源可能已经不再是一个稳定的视频源，应该对这一不稳定的视频源有所动作。 445

对于没有维持正确SCH关系的不稳定视频源（例如模拟VCR），当信号需要存储器寻址或下游处理时，应该生成合成的FIELD_0和FIELD_1输出（例如，通过将F输出信号除以2和4）。

❑ PAL场标识

场1和场5的开头可以通过监测外插的色同步的－U分量与同步之间的关系确定。如图8-16所示，在场1的开头，相对于同步，理想的相位为0°；在场5的开头，相对于同步，理想的相位为180°。无论是色同步消隐序列还是副载波相位，都可以用于区分场1和场3、场2和场4、场5和场7以及场6和场8。上一节中针对NTSC的所有考虑都适用于PAL。

9.2.9 视频信号类型自动检测

如果解码器可以自动地检测被解码视频信号的类型，并对自身进行自动配置，那么用户将不用猜测要处理视频信号的类型。这个信息可以通过状态信息传递给系统的其他部分。

如果解码器检测到至少16个连续帧每帧都少于575行，那么解码器可以假定视频信号是（M）NTSC或（M）PAL制式。首先，假定视频信号是（M）NTSC制式，因为这一制式非常流行。如果场时序和行时序保持锁定，但是解码器不能维持副载波锁定，那么视频信号可能是（M）PAL制式。在这种情况下，可以尝试（M）PAL操作，验证色同步时序。

如果解码器检测到至少16个连续帧每帧都大于575行，那么解码器可以假定视频信号是（B、D、G、H、I、N、N_C）PAL制式或者某种版本的SECAM制式。

首先，假定视频信号是（B、D、G、H、I、N）PAL制式。如果场时序和行时序保持锁定，但是解码器无法维持一个副载波锁，这可能意味着视频信号是（N_C）PAL制式或者SECAM制式。在这种情况下，可以尝试SECAM操作（因为它非常流行）。如果还是无法维持副载波锁，可以尝试（N_C）PAL操作。

如果解码器检测到一个它无法锁定的视频信号格式，那么解码器应该标识出来，这样用户就

可以知道。

注意在模拟VCR的特殊功能模式（例如快进）期间不能执行自动检测。如果解码器检测到一个特殊功能模式的操作，解码器应该禁用自动检测电路。只有输入视频信号丢失后检测视频信号时，才允许进行自动检测。

9.2.10　Y/C分离技术

编码器通常通过将亮度信号和色度信号简单地叠加在一起将它们组合起来；这样色度信号和高频亮度信号就会占用频谱的相同部分。结果，困难的是还要在解码器中分离它们。当解码信号时，一些亮度信息作为彩色信息解码（称之为串色），一些色度信息停留在亮度信号中（称之为串亮度）。由于数字解码器的稳定性能，可以采用比处理模拟解码器更复杂的分离技术。

446

在编辑场合中，串扰并不是什么好事。第一次解码的串扰分量又被编码，当下一次解码时，可能会产生新的或者额外的伪影。另外，当在解码信号中捕捉到静止帧时，边上的冻结残留副载波可能与随后编码过程的副载波冲突，致使彩色区域中的边闪烁。尽管解码器无法完全解决串扰问题，但是细粒度的Y/C分离可以使这一问题的影响降到最小。

如果解码器用在编辑环境中，那么对串亮度和串色的抑制要比被解码图像的外观更加重要。当图像被解码、处理、编码、再次解码时，各处理过程中相互之间的影响可能会引入实质性的伪影。最好是限制亮度带宽（以降低串亮度），生成较柔和的图像。另外，限制色度带宽小于1MHz，可以以色度清晰度为代价减少串色。

互补Y/C分离保留所有的输入信号。如果分离的色度信号和亮度信号再次叠加在一起，就再次生成原始的复合视频信号。

非互补Y/C分离引入一些不能复原的损失，如果分离的色度信号和亮度信号再次叠加在一起生成复合视频信号，就会使频谱之间产生间隙。该损失是由于使用了较窄带的滤波器降低串色和串亮度所致。因此，非互补滤波通常不适合于必须进行多次编码和解码的操作，因为随着解码操作次数的增加，频谱间隙也会持续增加。但是，为了最佳的观看效果，它确实可以用来调整亮度和色度响应。

1. 简单Y/C分离

对于所有这些实现，垂直色度解析度没有任何损失，但是也没有采用任何串色抑制。对于PAL制式，微分相位失真使行间的误差无法抑制，导致称之为Hanover条的垂直模式。

简单Y/C分离器的可注意到的伪影是垂直边上的彩色伪影。这些包括彩色振铃、彩色拖尾和高频率灰度级信息中的彩虹显示。

❑ 低通滤波和高通滤波

最基本的Y/C分离器假定低于一特定点的频率是亮度，高于这一点的是色度。简单Y/C分离器的一个例子如图9-34所示。

低于3.0MHz（NTSC）或者3.8MHz（PAL）的频率假定为亮度。高于这些的频率假定为色度。这种技术不仅会损失高频率的亮度信息，而且假定为色度信息的信息，还会导致串色。

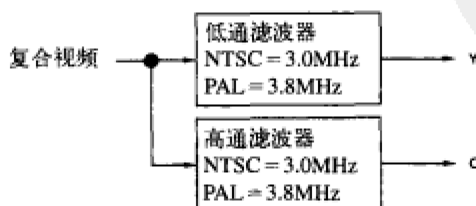


图9-34　典型的简单Y/C分离器

❑ 陷波滤波

尽管广播NTSC和PAL制式严格限制带宽，但是对于其他视频源来说情况并非如此。亮度信息可能位于6MHz、7MHz或者更高频率上。由于这一原因，图9-35中的设计通常更合适，因为它们允许高频亮度信息通过，产生较锐利的图像。

447

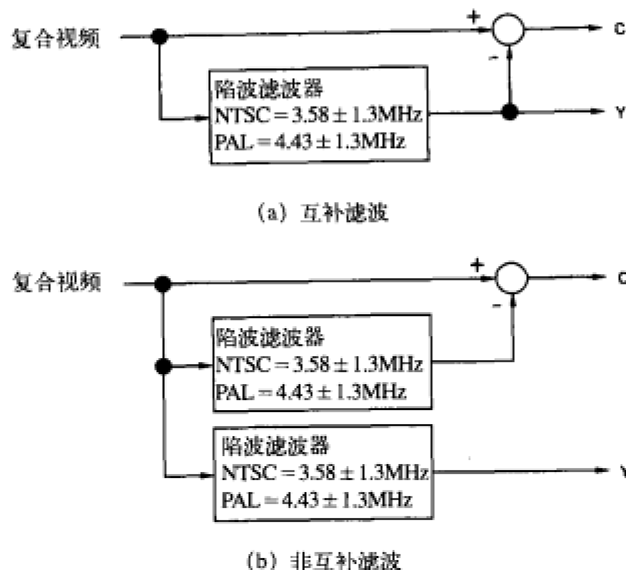


图9-35 典型的简单Y/C分离器

448

许多基于陷波滤波器的设计还在Y数据和色差数据路径上集成梳状滤波器，降低串色和串亮度伪影。但是，陷波滤波器仍然会限制整体的Y/C分离质量。

2. PAL制式注意事项

前面提到过，根据V分量是正常的还是倒置的，PAL制式使用正常和倒置的扫描行，以帮助校正由于微分相位失真引起的彩色偏移。

例如，微分相位失真可能导致正常扫描行上的绿色矢量角由图8-11中所示的理想241°滞后45°，导致矢量角为196°，有效地使结果颜色向黄色偏移。在倒置扫描行上，矢量角也将由图8-12中所示的理想120°滞后45°，导致矢量角为75°，有效地使结果颜色向青色偏移。

❑ PAL延迟行

图9-36由图8-12对U色轴180°削波，并将结果叠加在图8-11上生成，描述了相位误差的消除。两个误差（正常扫描行上的196°和倒置扫描行上的286°）的平均相位为241°，它是绿色的正确相位。由于这一原因，简单PAL解码器通常使用一个延迟行（或者行存储）来促进两个扫描行之间的平均。

在PAL制式Y/C分离器中使用延迟行有独特的问题。副载波基准在一个行周期内改变－90°（或者270°），V副载波在交变行上倒置。因此，行延迟的输入和输出之间有270°的相位差。如果我们希望在延迟行的输入和输出之间进行一个简单的加法或者减法，以恢复色度信息，那么相位差必须为0°或者180°，而且V的切换仍然需要。因此，我们希望寻找一种方法来调整行之间的副载波相位，补偿V的切换。

简单的电路，例如图9-37中所示的非互补Y/C分离器，使用了一个不是整行的延迟行（283.75个副载波周期），而不是284个副载波周期。这个微小的差别充当副载波频率的一个90°相位偏移。

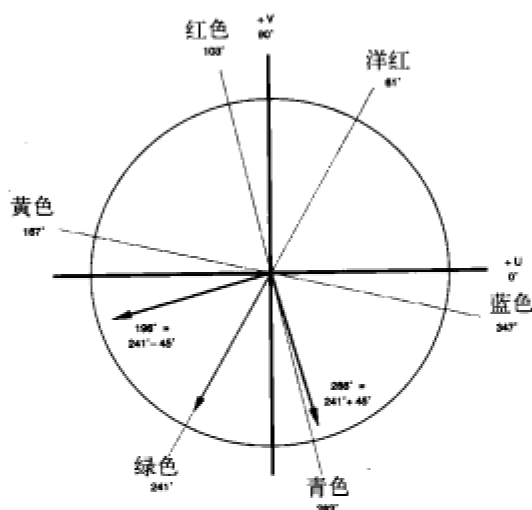


图9-36　PAL制式的相位误差校正

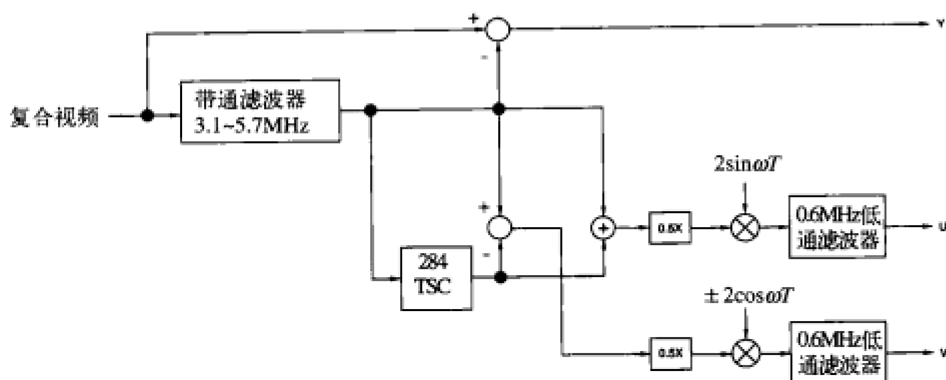


图9-37　简单延迟行PAL制式Y/C分离器

因为延迟中存在一个整数的副载波周期，所以位于第284个T_{SC}延迟行的输入和输出上的U副载波同相，而且它们可以简单地叠加在一起，恢复U副载波。由于V的切换，位于第284个T_{SC}延迟行的输入和输出上的V副载波180°异相，加法器将它们抵消。U解调器抑制所有的剩余高频垂直V分量。

由于V的切换，抽除第284个T_{SC}延迟行的输入和输出可以恢复V副载波，并且还可以清除U副载波。V解调器抑制所有的剩余高频垂直U分量。

因为通过第284个T_{SC}延迟行的相位偏移是频率的一个函数，所以副载波边带的相位偏移并不是正好90°，导致垂直色度转换上的色调误差。另外，因为色度向下偏移半行，所以色度和亮度没有垂直对齐。

❑ PAL修正器

尽管图9-37中电路的性能通常已经足够，但是第284个T_{SC}延迟行可以替换为一个行延迟后跟一个PAL修正器，如图9-38所示。PAL修正器提供一个90°的相位偏移和V副载波的倒置。PAL修正器的色度现在和行延迟输入同相，从而允许它们两个使用一个单独的加法器结合在一起，共享一个到解调器的普通路径。平均过程牺牲了一些垂直解析度，但是，Hanover条得到了抑制。

449

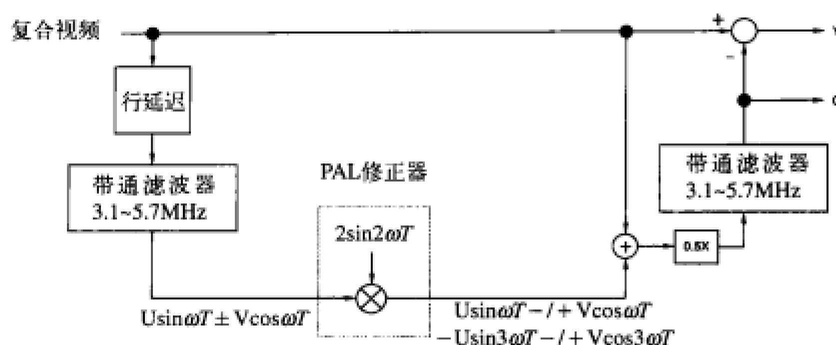


图9-38 使用PAL修正器的单行延迟PAL制式Y/C分离器

因为解调器输入的色度和复合视频同相，所以它可以用于清除复合信号中的色度，留下亮度。但是，由于色度向下偏移了半行，色度和亮度仍然没有垂直对齐。

PAL修正器生成一个中心位于两倍副载波频率处的亮度混叠。如果PAL修正器前面没有带通滤波器以及行之间的平均，那么将原始亮度分量和混叠亮度分量混合在一起将生成一个12.5Hz的拍频（beat frequency），在图像的高对比度区域可以注意到该拍频。

3. 2D梳状滤波

在前面的Y/C分离器中，高频的亮度信息被视为色度信息，而且还没有尝试着区分它们。因此，亮度信息被解释为色度信息（串色），传递给色度解调器，以恢复彩色信息。解调器不区分色度和高频亮度，从而在不应该有色彩的地方生成色彩。因此，生成偶发的显示伪影。

2D（或者场内）梳状滤波试图以降低垂直解析度为代价，提高色度和亮度的分离。梳状滤波器是因为使亮度和色度频率响应看上去像一把梳子而得名。理想情况下，这些频率响应应该与图8-4和图8-15中所示的隔行扫描的亮度和色度信号的梳子状的频率响应匹配。 450~451

现代的3行梳状滤波器通常使用2行延迟来存储视频信息的最后两行（使用这一方法在解码过程中有1行延迟）。使用多于2行的延迟通常导致过量的垂直滤波，降低垂直解析度。

❑ 两行延迟梳状滤波器

BBC对各种不同的梳状滤波实现（图9-39~图9-42）进行了广泛的研究（参考文献4），评价了每种实现的伪影和频率响应。每个梳状滤波器的垂直频率响应如图9-43所示。

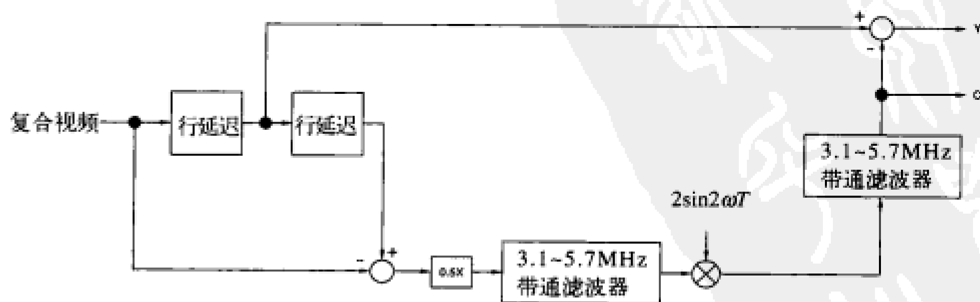


图9-39 2行Roe PAL制式Y/C分离器

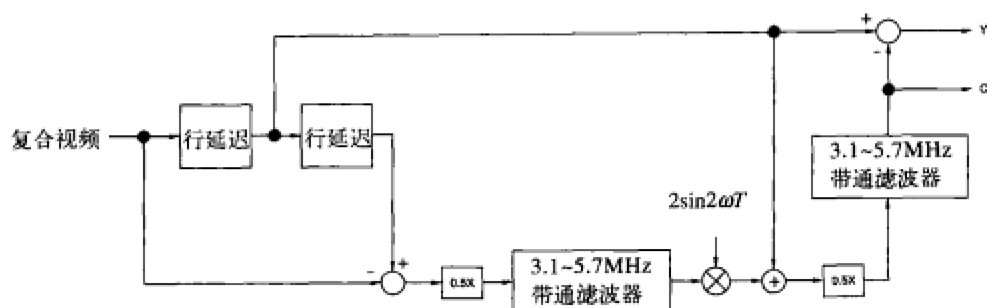


图9-40　2行－6dB Roe PAL制式Y/C分离器

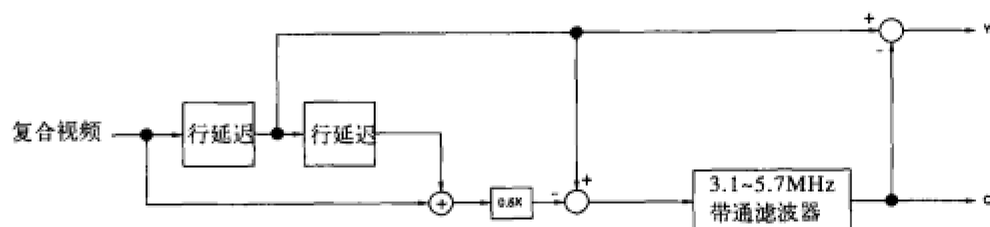


图9-41　2行余弦PAL制式Y/C分离器

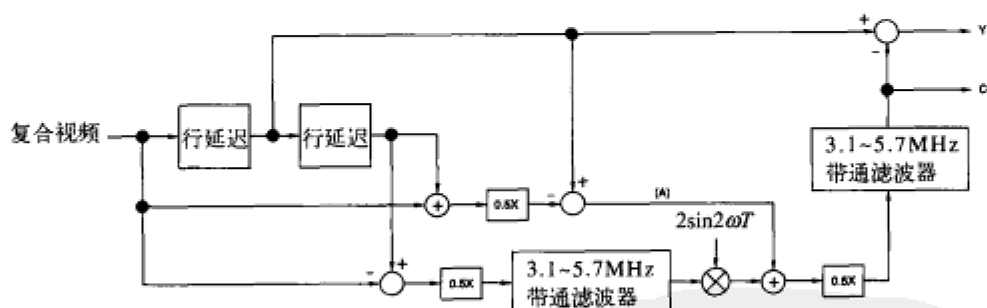


图9-42　2行Weston PAL制式Y/C分离器

在图9-39的梳状滤波器设计中，色度相位2行延迟后倒置。减法器清除了大部分的亮度，剩下双幅度、垂直滤波的色度。PAL修正器提供一个90°的相位偏移和PAL开关倒置的移除，以相位调整有1行延迟复合视频信号的色度。从复合信号中抽除色度后就剩下亮度。该设计的优点是色度和亮度垂直对齐。但是，这样会存在垂直解析度损失的问题，并且不抑制Hanover条。另外，在一些环境中，有可能由于PAL修正器生成的混叠亮度分量而生成双幅度的亮度。

图9-40中的梳状滤波器与图9-39中所示的梳状滤波器相似。但是，PAL修正器后面的色度以及1行延迟的复合视频信号叠加在一起生成双幅度的色度（因为它们的副载波同相）。再一次从复合信号中抽除色度后就剩下亮度。在该设计中，因为真实的和混叠的亮度信号都被平分，所以避免了亮度超出量程。这种设计中垂直解析度的损失较小，而且抑制Hanover条，但是以串色的增加为代价。

图9-41中的梳状滤波器设计的优点在于不用使用PAL修正器。因为色度相位2行延迟后倒置，将它们叠加在一起就会消除大部分的色度，剩下双幅度亮度。从1行延迟的复合视频信号中抽除该信号就生成色度。然后从1行延迟的复合视频信号中抽除色度就生成亮度（这样可以维持垂直亮度解析度）。因为行间的副载波相位没有对齐致使没有在加法器中消除色度信号，所以UV串扰

就表现为水平色度边上的12.5Hz闪烁。因为没有PAL修正器，所以没有亮度混叠或者亮度超标。

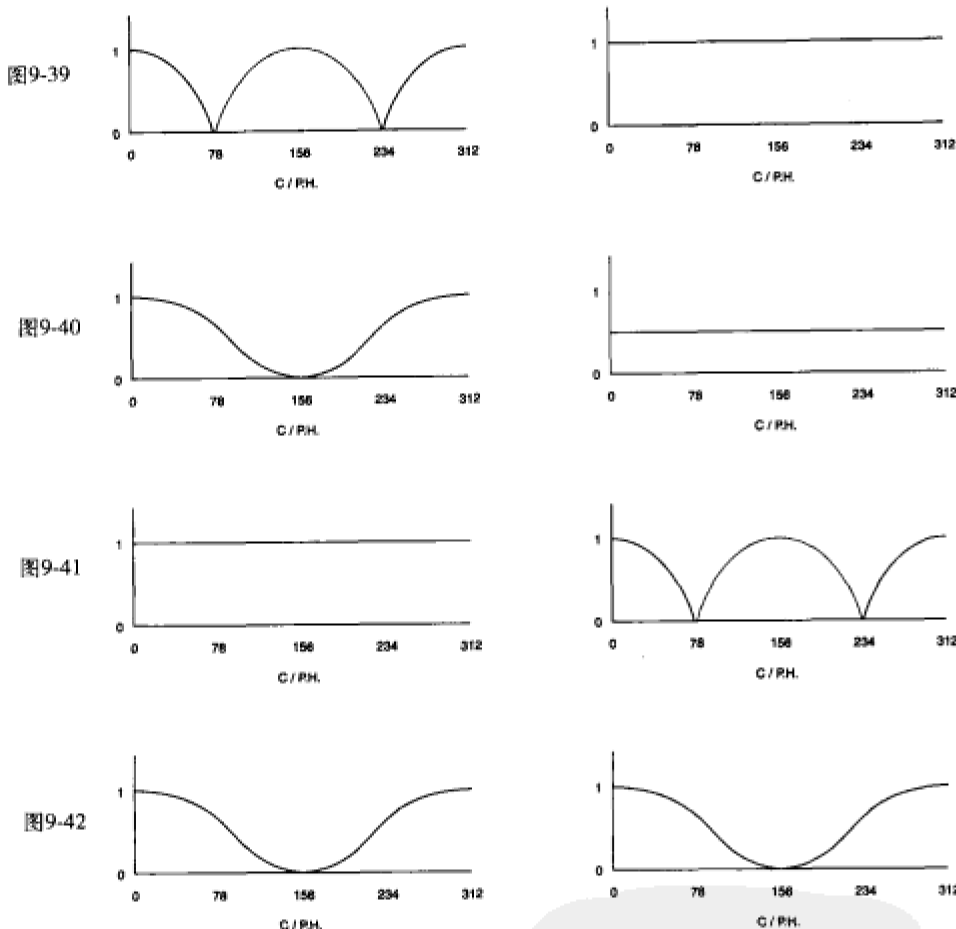


图9-43 图9-39~图9-42中梳状滤波器的垂直频率响应特性

图9-42中的梳状滤波器是图9-39和图9-41中梳状滤波器的组合。色度相位两行延迟后倒置。加法器消除大多数的色度，剩下双幅度亮度。从1行延迟复合视频信号中抽除该信号就生成色度信号（A）。在一个平行路径中，减法器消除大部分的亮度，剩下双幅度垂直滤波的色度。PAL修正器提供一个90°的相位偏移和PAL开关倒置的移除，以相位调整（A）色度信号。这些叠加在一起，生成双幅度的色度。然后从1行延迟的复合视频信号中抽除色度生成亮度。色度和亮度的垂直频率响应是图9-39和图9-41相应值的平均值。UV串扰与图9-41中的相似，但是幅度为图9-41中的一半。亮度混叠也是图9-39中的一半，而且抑制Hanover条。 452

以这些梳状滤波器设计为基础，BBC为普通电视设计了一种优化的梳状滤波器（图9-44）和标准变换（图9-45）。

对于PAL应用，最佳的亮度处理（图9-41）与最佳的色度处理（图9-40）结合在一起。两个设计之间的区别是色度恢复。对于标准变换（图9-45），色度信号正好是完整带宽的复合视频信号。标准变换使用垂直插值，有助于减少移动和高频垂直分量，包括串亮度和串色。因此，处理之后的垂直色度解析度通常将优于从普通电视的电路中所获得的垂直色度解析度。普通电视的电路（图9-44）以色度垂直解析度为代价、以减少串扰为目标恢复色度。

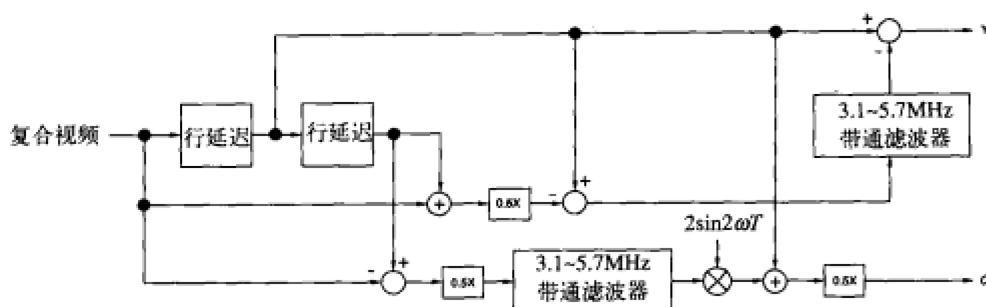


图9-44　为普通电视优化的2行延迟PAL制式Y/C分离器

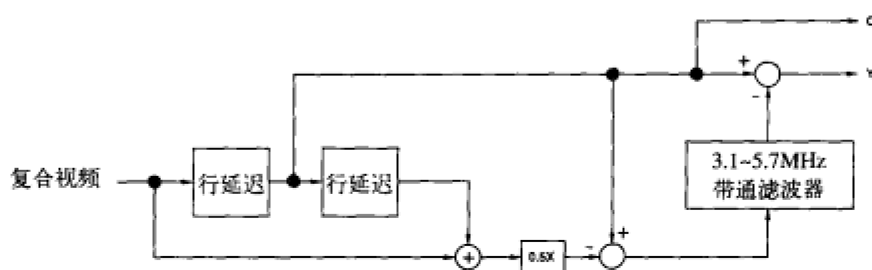


图9-45　为标准变换和视频处理优化的2行延迟PAL制式Y/C分离器

453～456

对于NTSC应用，梳状滤波器的设计相对简单。不用考虑副载波的切换，而且色度相位在每行上都是180°，而不是270°。另外，亮度频段和色度频段之间的分离要比PAL中的大，从而简化了分离要求。

在图9-46和图9-47中，因为副载波同相，所以加法器生成一个双幅度的复合视频信号。加法器的输出和1行延迟的复合视频信号之间有180°的副载波相位差，因此抽除这两个信号之后就消除了大部分的亮度，剩下双幅度的色度。

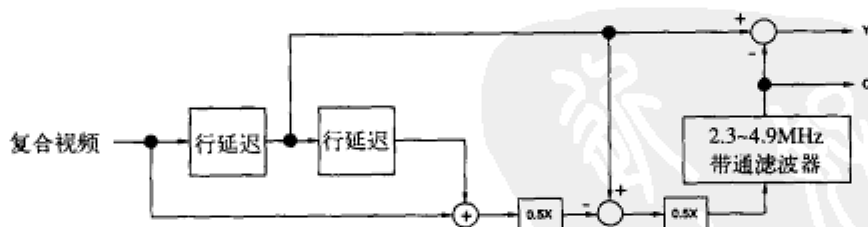


图9-46　普通电视的2行延迟NTSC制式Y/C分离器

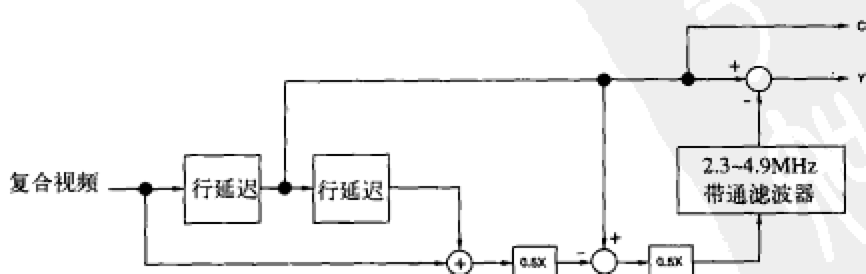


图9-47　标准变换和视频处理的2行延迟NTSC制式Y/C分离器

图9-46设计的主要缺点在于垂直色彩跳变上不对串亮度抑制。但是，这可以通过简单的低通

滤波所提高的亮度响应进行补偿。图9-47中处理色度的原因与图9-45中PAL制式的原因相同。

❑ 自适应梳状滤波

因为传统的梳状滤波器只对垂直对齐的样本进行处理，所以传统的梳状滤波器在对角线和垂直色彩变化方面仍然存在问题。

457

对于对角线，在标准梳状滤波之后，色度信息还包括邻接亮度值之间的差别，该差别也可以被解释为色度信息。该信息可以显示为串色伪影，例如扫描行边上出现的彩虹。

锐利垂直色彩跳变生成通常位于2个色彩转变之间的扫描行上的悬挂点（hanging dot）模式。在标准梳状滤波之后，亮度信息包含彩色副载波。彩色副载波的幅度由两个色彩之间的差别确定。因此，不同的色彩以不同的方式调制亮度光强度，在2个色彩之间的扫描行上生成一个点模式。为了消除这些悬挂点，有时候在梳状滤波器之后使用一个色度陷波滤波器。

自适应梳状滤波器试图通过处理一个3×3、5×5或者更大块的样本来解决这些问题。样本的值用来确定中心样本使用哪个Y/C分离算法。而Y/C分离算法多达32种，甚至更多。通过查看亮度的锐利垂直跳变，或者锐利彩色副载波相位变化，可以修改梳状滤波器的操作，以避免生成伪影。

由于组合行存储的代价，家用市场通常使用3行自适应梳状滤波，并进一步进行改进，也即3D运动自适应梳状滤波。

4. 3D梳状滤波

这个方法（也称为场间Y/C分离）使用当前场和前2场（NTSC）或前4场（PAL）的复合视频数据。这两个数据叠加在一起就可以消除色度（因为它是180°异相），剩下亮度。抽除这两个数据可以消除亮度，剩下色度。对于PAL，通过将图9-42中的行延迟替换为帧延迟可以获得一个适当的设计。

这一技术为静态图像提供了近乎完美的Y/C分离。但是，如果场间存在变化，所获得的Y/C分离结果之中就会包含错误。由于这一原因，通常不使用场间Y/C分离器，除非作为3D运动自适应梳状滤波器的部分时才使用。

❑ 3D运动自适应梳状滤波器

图9-48示范了一个典型的3D运动自适应梳状滤波器实现。在该实现中，静止区域使用3D（场间）梳状滤波器，而包含运动的图像区域使用2D（场内）梳状滤波器。运动检测器生成一个为0~1的值（K），允许2个梳状滤波器的亮度信号和色度信号按比例混合在一起。算法之间的硬切换通常是可见的。

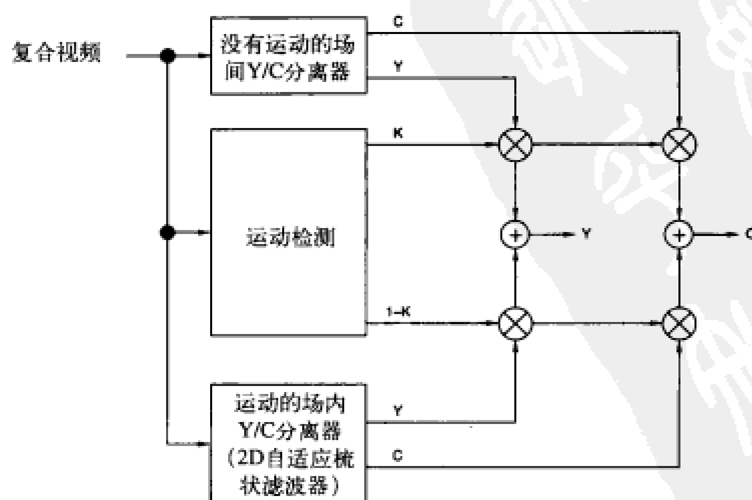


图9-48 3D运动自适应Y/C分离器

图9-49描述了一个简单的运动检测器框图。这里的思想就是在低频亮度信号中比较帧间的变化。在很大程度上，运动检测器的性能决定图像的质量。运动信号（K）通常被整流，通过平均水平和垂直方向的几个样本使运动信号平滑，并乘以一个增益因子，然后在使用之前进行削波。运动检测器容易犯的唯一错误就是在图像的静态区域使用2D梳状滤波器。

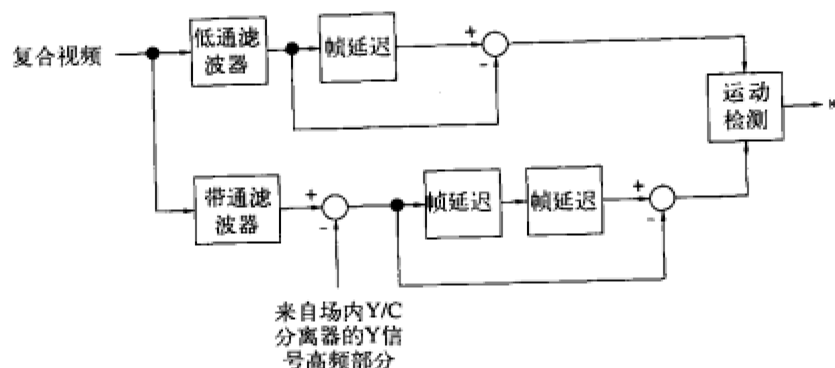


图9-49 NTSC制式的简单运动检测器框图

9.2.11 alpha通道支持

通过在NTSC/PAL解码器中集成一个额外的ADC，还可以数字化模拟alpha信号（也称为键），而且和视频数据一起流水，维持同步关系。这允许设计者改变解码器（可能拥有不同的流水线延迟），以适应具体的应用，而且不必考虑alpha通道的流水线延迟。alpha通常是线性的，模拟范围为0~100 IRE。alpha通道中不提供消隐基底或者同步信息。

458

9.2.12 解码器视频参数

为了指定NTSC/PAL解码器的相对质量，定义了许多工业级标准的视频参数。为了测量这些参数，NTSC/PAL解码器的输出（解码第8章中描述的各种不同视频测试信号）通过使用视频测试设备进行监测。表9-11示范了家用级和演播室级质量解码器的典型AC参数值以及个别参数的描述。

表9-11 NTSC和PAL解码器的典型AC视频参数

参数	家用级质量	演播室质量0	单位
微分相位	4	≤1	度
微分增益	4	≤1	%
亮度非线性特性	2	≤1	%
色调准确度	3	≤1	度
色饱和度准确度	3	≤1	%
SNR（遵循EIA/TIA RS-250-C）	48	>60	dB
色度－亮度串扰	<−40	<−50	dB
亮度－色度串扰	<−40	<−50	dB
H倾斜	<1	<1	%
V倾斜	<1	<1	%
Y/C采样偏离	<5	<2	ns
解调正交	90±2	90±0.5	度

有几个AC参数，例如短时间波形失真、组延迟和K因子，取决于模拟视频滤波器的质量，因此这里不对它们进行讨论。除了本节讨论的AC参数外，还有一些其他参数应该包含在解码器规范中，例如色同步捕捉与锁定频率范围，以及被解码YIQ或YUV视频信号的带宽。

另外还应该指定几个DC参数，如表9-12所示。尽管通常不指定同步锁相能力，但是除了时钟抖动，我们还是试图生成一列的同步锁相参数，如表9-13所示。

表9-12 NTSC和PAL解码器的典型DC视频参数

参　　数	(M) NTSC	(B、D、G、H、I) PAL	单　　位
同步输入幅度	40±20	43±22	IRE
色同步输入幅度	40±20	42.86±22	IRE
视频输入幅度（标称1V）	0.5~2.0	0.5~2.0	V

表9-13 NTSC和PAL解码器的典型同步锁相参数。这些参数假定视频信号的SNR≥30dB，并且优于表9-12所示的DC参数

参　　数	最　　小	最　　大	单　　位
同步锁定时间[1]		2	场
同步恢复时间[2]		2	场
短期同步锁范围[3]	±100		ns
长期同步锁范围[4]	±5		μs
校正之前连续丢失行同步脉冲数量	5		同步脉冲
场相关性[5]		±5	ns
短期副载波锁定范围[6]	±200		Hz
长期副载波锁定范围[7]	±500		Hz
副载波锁定时间[8]		10	行
副载波准确度		±2	度

注：1. 从同步锁相过程开始到实现场相关性规范的时间。
2. 从同步锁相损失到实现场相关性规范的时间。
3. 维持场相关性规范的范围。短期范围假定行时间在两个连续行之间的改变量缓慢。
4. 维持场相关性规范的范围。长期范围假定行时间在一场中的改变量缓慢。
5. 指示场样本准确度。对于使用VCO或者VCXO的同步锁相系统，这一规范与样本时钟抖动相同。
6. 副载波锁定时间和维持准确度规范的范围。短期时间假定副载波在2帧内改变量缓慢。
7. 副载波锁定时间和维护准确度规范的范围。长期时间假定副载波在24小时内改变量缓慢。
8. 在副载波瞬时180°相位偏移之后，上锁时间位于±2°之内。副载波频率的标称值为±500Hz。

1. 微分相位

微分相位失真，通常称为微分相位，指的是色度相位受亮度电平的影响有多大，换句话说，就是当亮度电平变化时色调偏移有多大。相位误差可能是正的也有可能是负的，因此微分相位表示为峰－峰测量值，以副载波相位的度数表示。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级解码器的微分相位参数接近于1°或者更小。

2. 微分增益

微分增益失真，通常称为微分增益，指的是色度增益受亮度电平的影响有多大，换句话说，

就是当亮度电平变化时色饱和度偏移有多大。信号有可能衰减也有可能放大，因此微分增益表示为两个电平之间的最大幅度变化，表示为最大色度幅度的百分比。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级解码器的微分增益参数接近于1%或者更小。

3. 亮度非线性特性

亮度非线性特性也称为微分亮度和亮度非线性失真，表示亮度增益受亮度电平的影响有多大，换句话说，就是解码的亮度电平和理想亮度电平之间的非线性关系。

459～461

通过使用一个非调制的5级或者10级阶梯测试信号，或者复合测试信息的调制5级部分，将最大级和最小级之间的差，表示为最大级的一个百分比，用来表示亮度非线性特性。尽管这个参数包含在微分增益和微分相位参数中，但是传统上这个参数单独指定。

4. 色度非线性相位失真

色度非线性相位失真指的是色度相位（色调）受色度幅度（色饱和度）的影响有多大，换句话说，就是当饱和度变化时色调偏移有多大。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，可以测量每个色度包的解码器输出。最大色调测量值和最小测量值之间的差就是峰－峰值。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

5. 色度非线性增益失真

色度非线性增益失真指的是色度增益受色度幅度（色饱和度）的影响有多大，换句话说，就是指已解码的色度幅度电平和理想色度幅度电平之间的非线性关系，这通常看作高饱和度色度信号的衰减。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，对解码器进行调整，从而使中间色度包（40 IRE）被正确解码。另2个解码色度包的幅度的测量值和标称值之间的最大差表示色度非线性增益失真，用IRE表示，或者表示为最坏包标称幅度的百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

6. 色度－亮度互调

色度－亮度互调通常称为交叉调制，表示亮度电平受色度的影响有多大。这可能是高饱和度色度电平削波或正交失真的结果，而且由于色饱和度的改变还可以表现为一个不规则的亮度变化。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，用解码的50 IRE亮度电平和被解码亮度电平之间的最大差来表示色度－亮度互调，色度－亮度互调可以用IRE表示，或者表示为一个百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

7. 色调准确度

462～464

色调准确度表示被解码的色调值与理想的色调值之间有多紧密。相位误差可能是正的也可能是负的，所以色调准确度是最坏情况下的正测量值和最坏情况下的负测量值（相对于标称值）之间的差，以副载波相位的度数表示。这个参数使用EIA或EBU的75%彩条作为测试信号进行测量。

8. 色饱和度准确度

色饱和度准确度表示被解码的饱和度值与理想的饱和度值之间有多紧密，该参数使用EIA或EBU的75%彩条作为测试信号。该测量值可能是增益也有可能是衰减，所以色饱和度准确度是最坏情况下的增益测量值和最坏情况下的衰减测量值（相对于标称值）之间的差，表示为标称值的

百分比。

9. H倾斜

H倾斜也称行倾斜和行时间失真，导致行频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像左边和右边之间亮度的变化。对于数字解码器，H倾斜主要是模拟输入滤波器和传输媒介的伪影。H倾斜使用一个线条（例如NTC-7 NTSC复合测试信号中的线条）来测量，并且测量倾斜的峰－峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的第一个和最后一个微秒。

10. V倾斜

V倾斜也称为场倾斜和场时间失真，导致场频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像的顶边和底边之间亮度的变化。对于数字解码器，V倾斜主要是模拟输入滤波器和传输媒介的伪影。V倾斜使用18μs、100 IRE的白条（位于场中心的130行扫描行的中心）或者场方波测量。所测量的是倾斜的峰－峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的前3行和最后3行。

参考文献

1. Benson, K. Blair, 1986, *Television Engineering Handbook*, McGraw-Hill, Inc.
2. Clarke, C.K.P., 1986, *Colour encoding and decoding techniques for line-locked sampled PAL and NTSC television signals*, BBC Research Department Report BBC RD1986/2.
3. Clarke, C.K.P., 1982, *Digital Standards Conversion: comparison of colour decoding methods*, BBC Research Department Report BBC RD1982/6.
4. Clarke, C.K.P., 1982, *High quality decoding for PAL inputs to digital YUV studios*, BBC Research Department Report BBC RD1982/12.
5. Clarke, C.K.P., 1988, *PAL Decoding: Multidimensional filter design for chrominanceluminance separation*, BBC Research Department Report BBC RD1988/11.
6. Drewery, J.O., 1996, *Advanced PAL Decoding: Exploration of Some Adaptive Techniques*, BBC Research Department Report BBC RD1996/1.
7. ITU-R BT.470-6, 1998, *Conventional Television Systems*.
8. *NTSC Video Measurements*, Tektronix, Inc., 1997.
9. Perlman, Stuart S. et. al., *An Adaptive Luma-Chroma Separator Circuit for PAL and NTSC TV Signals*, International Conference on Consumer Electronics, Digest of Technical Papers, June 6-8, 1990.
10. Sandbank, C. P., *Digital Television*, John Wiley & Sons, Ltd., 1990.
11. SMPTE 170M-1999, *Television-Composite Analog Video Signal-NTSC for Studio Applications*.
12. *Television Measurements, NTSC Systems*, Tektronix, Inc., 1998.
13. *Television Measurements, PAL Systems*, Tektronix, Inc., 1990.

第10章 H.261和H.263

目前有多个视频会议标准，如表10-1所示。图10-1至图10-3举例说明了几个常见的视频会议系统的结构图。

表10-1 视频会议标准系列

	H.310	H.320	H.321	H.322	H.323	H.324	3G-324M
网络	宽带ISDN ATM LAN	窄带交换数据ISDN	宽带ISDN ATM LAN	保证带宽分组交换网络	不保证带宽分组交换网络（以太网）	PSTN或POTS	移动网络
视频编解码器	MPEG-2 H.261	H.261 H.263 H.264	H.261 H.263	H.261 H.263	H.261 H.263 H.264	H.261 H.263	MPEG-4.2
音频编解码器	MPEG-2 G.711 G.722 G.728	G.711 G.722 G.722.1 G.728	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.722.1 G.723.1 G.728 G.729	G.723	G.722.2 G.723.1
复用	H.222.0 H.222.1	H.221	H.221	H.221	H.225.0	H.223	H.223A/B
控制	H.245	H.231 H.242 H.243	H.242	H.230 H.242	H.245	H.245	H.245
多点		H.231	H.231	H.231	H.323		
数据	T.120	H.239 T.120	T.120	T.120	H.239 T.120	T.120	T.120
通信接口	AAL I.363 AJM I.361 PHY I.432	I.400	AAL I.363 AJM I.361 PHY I.400	I.400 和 TCP/IP	TCP/IP	V.34 调制解调器	移动通信

10.1 H.261

ITU-T H.261是第一个为视频会议系统而开发的视频压缩和解压标准。H.261最初是为在窄带综合业务数字网（N-ISDN）上开展速率为$p \times 64$ kbit/s的可视电话和视频会议等业务而制定的，其中$p=1\sim30$。现在，H.261已经成为了所有视频会议标准的最低需求，如表10-1所示。

图10-4给出了一个典型的H.261编码器结构图。视频编码器提供了一个与其他信号（如控制信号、音频信号等）复用的独立数字视频比特流。视频解码器执行与编码器相反的过程。

如图3-7所示，H.261视频数据采用4：2：0的YCbCr格式。表10-2中给出了其主要的规格说明。H.261最大的图像速率是受限的，即在两个传送图像之间可能有0、1、2或3个非传送图像。

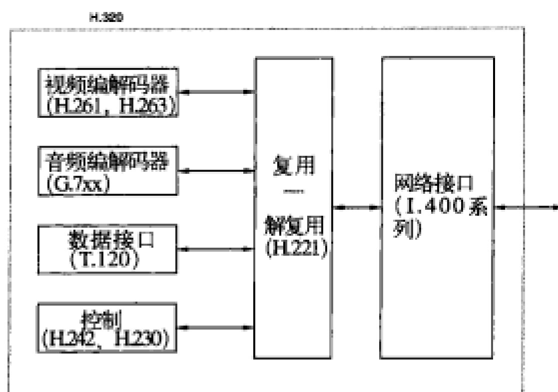


图10-1 典型的H.320系统

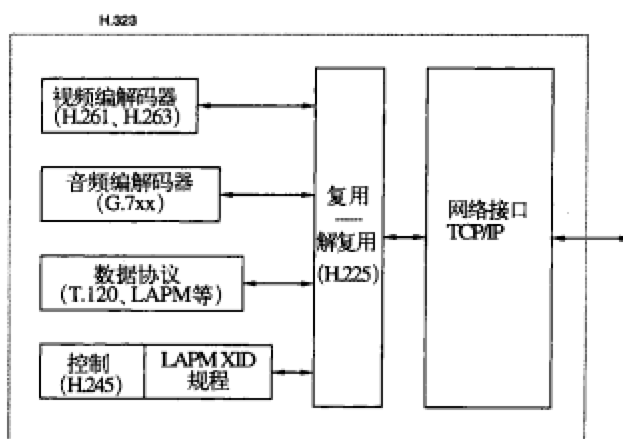


图10-2 典型的H.323系统

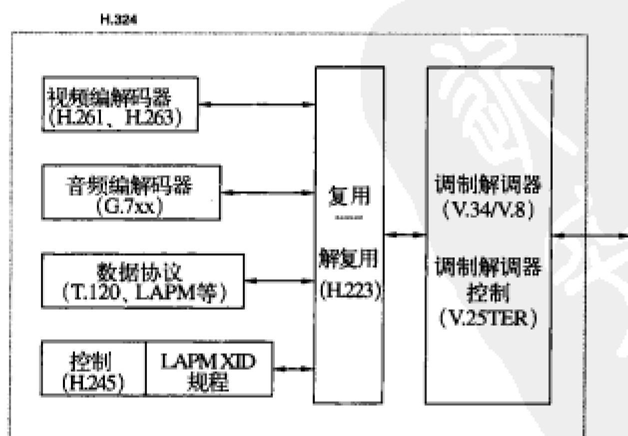


图10-3 典型的H.324系统

表10-2　H.261的YCbCr参数

参　　数	CIF	QCIF
有效分辨率	352 × 288	176 × 144
帧率	29.97Hz	
YCbCr采样结构	4：2：0	
YCbCr编码形式	归一化量化DCT，每采样点8位表示	

H.261支持的两种图像（或帧）类型如下。

- 内码帧（Intra）或I帧：没有使用参考帧进行预测的帧。
- 预测帧（Inter）或P帧：基于前一帧预测得到的帧。

10.1.1　视频编码层

如图10-4所示，编码器基本的功能有预测、块变换和量化。

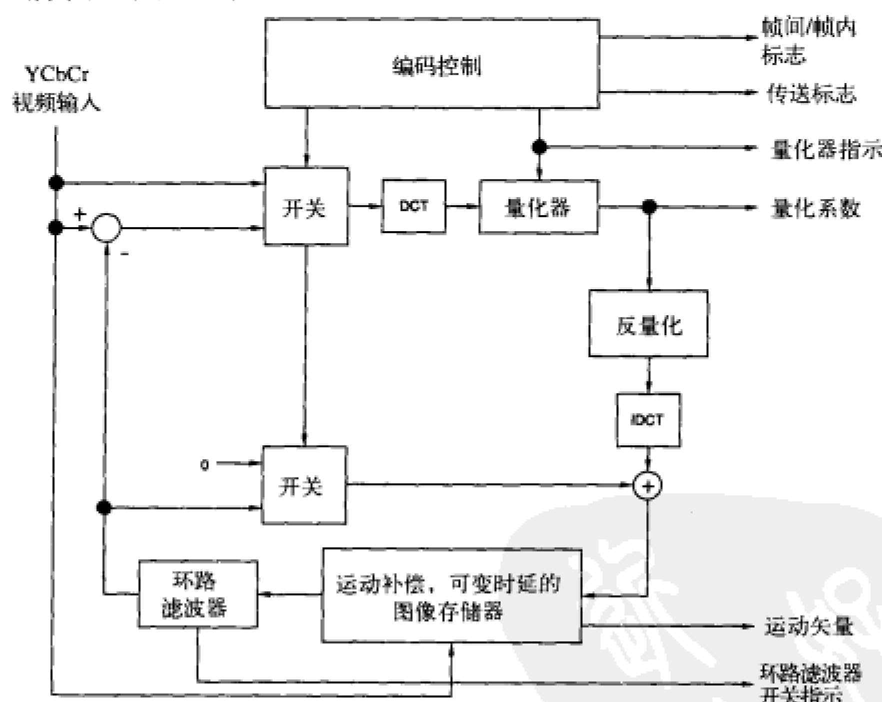


图10-4　典型的H.261编码器

预测误差（帧间模式）或者输入图像（帧内模式）被细分为多个8像素 × 8行的块，这些块被分成传送的和非传送的。如图10-5所示，4个亮度块（Y）和两个空间上对应的色差块（Cb和Cr）组成了一个16像素 × 16行的宏块（macroblock）。

H.261标准中没有给出模式选择和块传送的准则，这些准则可以作为编码策略的一部分动态地变化。编码器首先对传送块进行变换，并对变换后的系数进行量化，然后对量化后的系数进行可变长编码。

466

1. 预测

通常，预测是在图像之间（inter-picture）进行，可以加入运动补偿和空域滤波器。采用预测的编码模式称为帧间编码（inter），而不采用预测的编码模式则称为帧内编码（intra）。

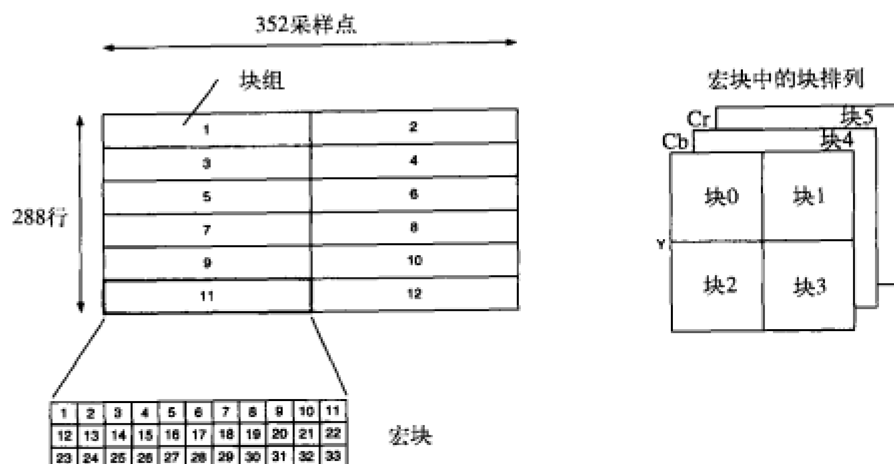


图10-5 H.261块组、宏块和块排列

2. 运动补偿

运动补偿在编码器中是可选的。解码器必须支持每个宏块拥有一个运动矢量。运动矢量是受限的，即所有参考像素都必须在已编码的图像区域内。

运动矢量的水平和垂直分量必须为不超过±15的整数值。宏块中所有4个Y块的运动矢量与宏块的运动矢量是相同的，而Cb和Cr块的运动矢量则为宏块运动矢量值的一半。

运动矢量的水平或垂直分量如果为正值，则表示预测是基于前一帧图像中位于被预测像素右方或者下方的像素。

3. 环路滤波器

可以采用二维空域滤波器来改进预测过程，该滤波器对8×8预测块内的像素进行操作。

滤波器可以分成水平和垂直函数，两者都是非递归的，且系数为0.25、0.5、0.25。特别地，在块边缘处，可能会有某一个抽头（tap）落入块区域之外，此时则将滤波器的系数改变为0、1、0。

467 ≀ 469

对于一个宏块中的所有6个块，滤波器的开与关将根据宏块的类型来决定。

4. 离散余弦变换和反离散余弦变换

在编码器中，传送的块首先进行8×8的离散余弦变换（DCT）处理。反离散余弦变换（IDCT）的输出结果削波之后的范围在−256和255之间，采用9位编码表示。

H.261标准也没有定义计算变换的过程，但是反变换必须满足特定的误差容限。

5. 量化

在一个宏块中，除了intra-DC外，所有系数都采用相同的量化器。intra-DC通常采用步长为8且无死区的线性量化器量化。而其他系数则采用中心死区为0，步长为2，4，6，…，62之间任意值的线性量化器量化。

6. 重构图像削波

为了防止编码器和解码器环路中由于变换系数振幅的量化失真可能导致的运算溢出，H.261标准中采用了削波函数。由预测值和预测误差相加形成重构图像，削波函数用于对重构图像进行削波处理。削波器将像素值限定为小于等于0或者大于等于255。

7. 编码控制

尽管没有作为H.261标准的一部分而包含在其中，但是H.261标准还是采用了多个参数来控制编码视频数据的产生速率。这些参数包括：编码预处理过程、量化器、块重要性准则和时域亚采

470–471

样等。时域亚采样是通过丢弃完整的图像来实现的。

8. 强制更新

为了防止IDCT变换的误差累积或者其他一些差错的影响，源编码采取了一种强制更新的方法。它规定每个宏块至少每传送132次就要以帧内模式传送一次（更新一次），以阻止编码中的误差累积。

10.1.2　视频比特流

除非特别声明，视频比特流中最高有效位MSB（Most Significant Bit）总是最先传送。MSB位于编码表的第1位，即最左边一位。同样，除非特别声明，所有的备用位均设置为1。

H.261的码流结构分为四个层次，从上到下依次是：

图像层（Picture）

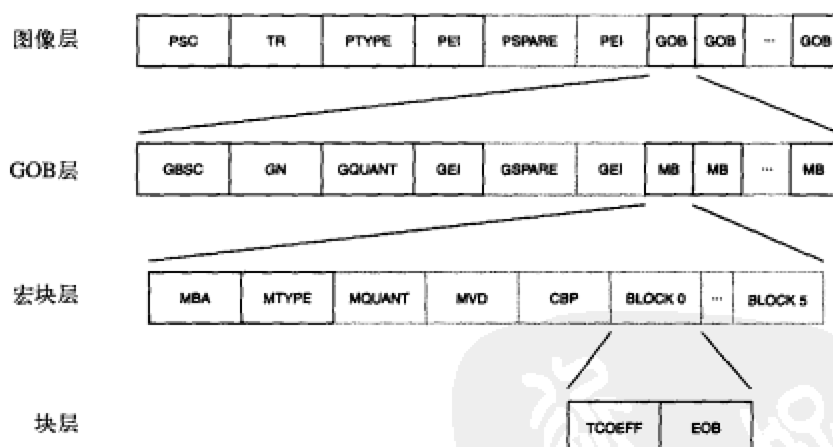
块组层GOB（Group of Blocks）

宏块层MB（Macroblock）

块层（Block）

1. 图像层

如图10-6所示，图像数据由图像头和块组（GOB）数据组成。被丢弃图像的图像头将不需要传送。



472

图10-6　H.261视频比特流层次结构

- **图像起始码（PSC）**

PSC（Picture Start Code）长度为20位，值为0000 0000 0000 0001 0000。

- **时间参考（TR）**

TR（Temporal Reference）是一个5位的二进制数，表示32种可能值。TR的生成是通过将前一图像中的TR值递增1，然后再加上非传送图像（在29.97Hz频率下）数目而得到的。运算仅在5个最低有效位LSB（Last Significant Bit）上进行。

- **类型信息（PTYPE）**

PTYPE（Type Information）采用6位编码来表示有关图像的信息：

位1表示分割屏幕指示器

0＝关，1＝开

位2表示文档相机指示器

0＝关，1＝开
位3表示冻结图像解除
0＝关，1＝开
位4表示源格式
0＝QCIF，1＝CIF
位5表示可选的静止图像模式
0＝开，1＝关
位6表示备用

❑ 额外插入信息（PEI）

PEI（Extra Insertion Information）采用1位表示，当设置为1时表示接下来将出现可选数据字段。

❑ 备用信息（PSPARE）

如果PEI设置为1，在PEI之后将紧随一个9位的额外插入信息，其中8位为PSPARE（Spare Information），第9位为PEI位，表示是否有另一个9位的额外插入信息在之后，如此反复。

2. 块组层

每一个图像都被分成了多个块组（GOB），CIF图像由12个GOB组成，而QCIF图像则由3个GOB组成（见图10-5）。GOB由176样本×48行的Y块和对应的88×24的Cb和Cr块数据组成。

如图10-6所示，GOB数据由GOB头和宏块数据组成。在图10-5中，在CIF或者QCIF中顺序编号的图像起始码之间，每一个GOB头将被传送一次，即使GOB中没有宏块数据也需传送。

❑ 块组起始码（GBSC）

GBSC（Group of Blocks Start Code）长度为16位，值为0000 0000 0000 0001。

❑ 组编号（GN）

GN（Group Number）是一个4位二进制值，表示块组的位置。这4位是图10-5中编号的二进制表示，编号13、14、15保留以备将来使用。

❑ 量化器信息（GQUANT）

GQUANT（Quantizer Information）是一个5位二进制值，表示块组使用的量化器，GQUANT一直有效，直到被接下来的MQUANT所取代。GQUANT取值为1~31。

❑ 额外插入信息（GEI）

GEI（Extra Insertion Information）用1位表示，当设置为1时，表示接下来将出现可选数据字段。

❑ GSPARE

如果GEI设置为1，则在GEI之后将紧随一个9位的额外插入信息，其中8位为GSPARE，第9位为GEI位，表示在之后是否有另一个9位的额外插入信息，如此反复。

3. 宏块层

如图10-5所示，一个GOB被分成33个宏块。每一个宏块由一个16样本×16行的Y块和对应的8×8的Cb块和Cr块组成。

宏块的数据由宏块头和块数据组成（见图10-6）。

❑ 宏块地址（MBA）

MBA（Macroblock Address）是一个变长的码字，表示宏块在块组中的位置。图10-5中给出了宏块的传送顺序。对于块组GOB中的第一个宏块，其MBA是图10-5中的绝对地址，对于接下来的宏块，它们的MBA是其绝对地址与最近传送的宏块的MBA之间的差值。表10-3中给出了MBA的码表。

表10-3 H.261中MBA的可变长编码表

MBA	编码			MBA	编码			
1	1			17	0000	0101	10	
2	011			18	0000	0101	01	
3	010			19	0000	0101	00	
4	0011			20	0000	0100	11	
5	0010			21	0000	0100	10	
6	0001	1		22	0000	0100	011	
7	0001	0		23	0000	0100	010	
8	0000	111		24	0000	0100	001	
9	0000	110		25	0000	0100	000	
10	0000	1011		26	0000	0011	111	
11	0000	1010		27	0000	0011	110	
12	0000	1001		28	0000	0011	101	
13	0000	1000		29	0000	0011	100	
14	0000	0111		30	0000	0011	011	
15	0000	0110		31	0000	0011	010	
16	0000	0101	11	32	0000	0011	001	
				33	0000	0011	000	
				MBA填充	0000	0001	111	
				起始码	0000	0000	0000	0001

对于在紧接在GOB头之后或者已编码的宏块之后的填充位（称为MBA填充），也可以有一个码字。这个码字在解码器中将被丢弃。

表10-3中同时也给出了起始码的码字。MBA通常包含在被传送的宏块中。当宏块中没有包含图像的MBA信息时，该宏块将不会被传送。

❑ 类型信息（MTYPE）

MTYPE（Type Information）是一个变长的码字，它包含有关于宏块和宏块中数据元素的信息。表10-4中列出了宏块类型、包含的元素和变长的码字。MTYPE通常也包含在被传送的宏块中。

表10-4 H.261中MTYPE的可变长编码表

预测模式	MQUANT	MVD	CBP	TCOEF	编码		
intra				×	0001		
intra	×			×	0000	001	
inter			×	×	1		
inter	×		×	×	0000	1	
inter＋MC		×			0000	0000	1
inter＋MC		×	×	×	0000	0001	
inter＋MC	×	×	×	×	0000	0000	01
inter＋MC＋FIL		×			001		
inter＋MC＋FIL		×	×	×	01		
inter＋MC＋FIL	×	×	×	×	0000	01	

❑ 量化器（MQUANT）

只有当MTYPE中指示MQUANT存在时，MQUANT（Quantizer）才会出现。MQUANT是一

个5位的码字，它表示块组中当前和接下来的块要采用的量化器，MQUANT一直有效，直到被下一个MQUANT所取代。MQUANT的码字与GQUANT的码字是一致的。

❑ 运动矢量数据（MVD）

正如MTYPE所指示的那样，所有的运动补偿（MC）宏块都包含MVD（Motion Vector Data）。MVD是通过将当前宏块矢量减去前一个宏块矢量所获得的。当出现下列情况时，前一个宏块的矢量被视为0：

（1）估计宏块1、12和23的MVD时；

（2）当前宏块MBA与前一宏块的MBA差值小于1时；

（3）前一宏块的MTYPE不是运动补偿时。

运动矢量数据由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-5中给出了这些变长码。

表10-5 H.261中MVD的可变长编码表

矢量差分	编码			矢量差分	编码		
–16和16	0000	0011	001	1	010		
–15和17	0000	0011	011	2和–30	0010		
–14和18	0000	0011	101	3和–29	0001	0	
–13和19	0000	0011	111	4和–28	0000	110	
–12和20	0000	0100	001	5和–27	0000	1010	
–11和21	0000	0100	011	6和–26	0000	1000	
–10和22	0000	0100	11	7和–25	0000	0110	
–9和23	0000	0101	01	8和–24	0000	0101	10
–8和24	0000	0101	11	9和–23	0000	0101	00
–7和25	0000	0111		10和–22	0000	0100	10
–6和26	0000	1001		11和–21	0000	0100	010
–5和27	0000	1011		12和–20	0000	0100	000
–4和28	0000	111		13和–19	0000	0011	110
–3和29	0001	1		14和–18	0000	0011	100
–2和30	0011			15和–17	0000	0011	010
–1	011						
0	1						

❑ 编码块模式（CBP）

如果MTYPE指示CBP存在，将出现一个变长的CBP（Coded Block Pattern）。CBP表示宏块中哪一个块至少传送了一个变换系数。模式编码如下所示：

$P_0P_1P_2P_3P_4P_5$

对于块[n]，其任意系数出现，则$P_n=1$，否则$P_n=0$。图10-5中给出了块编号方式（十进制）。表10-6中给出了CBP编码的码字。 474

表10-6 H.261中CBP的可变长编码表

CBP	编码		CBP	编码	
60	111		62	0100	0
4	1101		24	0011	11
8	1100		36	0011	10

（续）

CBP	编 码		CBP	编 码		
16	1011		3	0011	01	
32	1010		63	0011	00	
12	1001	1	5	0010	111	
48	1001	0	9	0010	110	
20	1000	1	17	0010	101	
40	1000	0	33	0010	100	
28	0111	1	6	0010	011	
44	0111	0	10	0010	010	
52	0110	1	18	0010	001	
56	0110	0	34	0010	000	
1	0101	1	7	0001	1111	
61	0101	0	11	0001	1110	
2	0100	1	19	0001	1101	
35	0001	1100	38	0000	1100	
13	0001	1011	29	0000	1011	
49	0001	1010	45	0000	1010	
21	0001	1001	53	0000	1001	
41	0001	1000	57	0000	1000	
14	0001	0111	30	0000	0111	
50	0001	0110	46	0000	0110	
22	0001	0101	54	0000	0101	
42	0001	0100	58	0000	0100	
15	0001	0011	31	0000	0011	1
51	0001	0010	47	0000	0011	0
23	0001	0001	55	0000	0010	1
43	0001	0000	59	0000	0010	0
25	0000	1111	27	0000	0001	1
37	0000	1110	39	0000	0001	0
26	0000	1101				

4. 块层

一个宏块由4个Y块和一个Cb块、一个Cr块组成（参见图10-5）。

如图10-6所示，一个8样本×8行的块的数据由变换系数的码字和块结束（EOB，end of block）标识组成。图10-5中给出了块传送的顺序。

❑ 变换系数（TCOEFF）

475～476

当MTYPE指示为intra时，宏块中的所有6个块的变换系数数据都将出现。否则，MTYPE和CBP将联合指出哪些块有系数数据传送。量化的DCT系数的传送顺序如图7-57所示。

如表10-7所示，大多数常见的连续零值（RUN）和紧接之后的非零值（LEVEL）的组合采用变长编码方式编码。因为CBP指示了哪些块没有系数数据，因此EOB不可能出现在第一个系数中。最后一位s表示非零值（LEVEL）的符号：0表示为正，1表示为负。

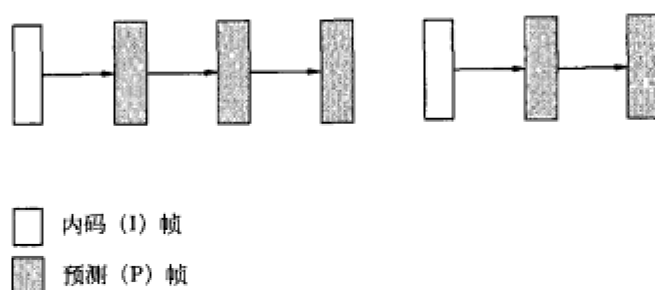


图10-7 典型的H.261解码序列

（RUN，LEVEL）的其他组合采用一个20位的码字编码：其中6位表示转义（ESC），6位表示RUN，8位表示LEVEL，如表10-8所示。

变长编码采用两个编码表：一个用于inter、inter＋MC和inter＋MC＋FIL块中的第一个传送的LEVEL；另一个用于除intra块中的第一个LEVEL之外的所有其他LEVEL，而对于intra块中的第一个LEVEL，则采用一个8位的定长编码。

除intra-DC外，所有的系数都有一个重构非零值（reconstruction level，REC），取值范围为–2048~2047。重构非零值可通过下面的方程获得，其结果已经经过削波处理。QUANT取值从1到31，QUANT一直有效，直到被另一个GQUANT或者MQUANT所取代。

QUANT＝奇数：

　若 LEVEL>0

　　REC＝QUANT × (2 × LEVEL＋1)

　若 LEVEL<0

　　REC＝QUANT × (2 × LEVEL-1)

QUANT＝偶数：

　若 LEVEL>0

　　REC＝(QUANT × (2 × LEVEL＋1))-1

　若 LEVEL<0

　　REC＝(QUANT × (2 × LEVEL-1))＋1

　若 LEVEL＝0

　　REC＝0

表10-7 H.261中TCOEFF的可变长编码表

RUN	LEVEL	编码			
EOB		10			
0	1	1s	如果第一个系数在块中*		
0	1	11s	如果第一个系数不在块中		
0	2	0100	s		
0	3	0010	1s		
0	4	0000	110s		
0	5	0010	0110	s	
0	6	0010	0001	s	
0	7	0000	0010	10s	
0	8	0000	0001	1101	s

（续）

RUN	LEVEL	编　　码			
0	9	0000	0001	1000	s
0	10	0000	0001	0011	s
0	11	0000	0001	0000	s
0	12	0000	0000	1101	0s
0	13	0000	0000	1100	1s
0	14	0000	0000	1100	0s
0	15	0000	0000	1011	1s
1	1	011s			
1	2	0001	10s		
1	3	0010	0101	s	
1	4	0000	0011	00s	
1	5	0000	0001	1011	s
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
2	1	0101	s		
2	2	0000	100s		
2	3	0000	0010	11s	
2	4	0000	0001	0100	s
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0100	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0011	0s		
4	2	0000	0011	11s	
4	3	0000	0001	0010	s
5	1	0001	11s		
5	2	0000	0010	01s	
5	3	0000	0000	1001	0s
6	1	0001	01s		
6	2	0000	0001	1110	s
7	1	0001	00s		
7	2	0000	0001	0101	s
8	1	0000	111s		
8	2	0000	0001	0001	s
9	1	0000	101s		
9	2	0000	0000	1000	1s
10	1	0010	0111	s	
10	2	0000	0000	1000	0s
11	1	0010	0011	s	
12	1	0010	0010	s	
13	1	0010	0000	s	
14	1	0000	0011	10s	
15	1	0000	0011	01s	

（续）

RUN	LEVEL	编码			
16	1	0000	0010	00s	
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
ESC		0000	01		

注：在intra宏块中不采用。

对于intra-DC块，第一个系数通常为采用步长为8且无死区的量化器量化的变换值，其结果为一个8位编码值*n*。黑色（black）的编码值为0001 0000（16），白色（white）的编码值为1110 1011（235）。变换值1024的编码值为1111 1111。编码值0000 0000和1000 0000没有使用。除了255以外，所有编码值的解码值都为8*n*，而255的解码值为重构变换值1024。

477
~
478

10.1.3 静止图像传送

H.261允许传送分辨率为当前所选视频格式分辨率4倍的静止图像。如果视频格式为QCIF，可以传送CIF分辨率的静止图像；如果视频格式为CIF，可以传送分辨率为704×576的静止图像。

表10-8 H.261 RUN和LEVEL编码

RUN	编码	LEVEL	编码
0	0000 00	−128	禁止
1	0000 01	−127	1000 0001
⋮	⋮	⋮	⋮
63	111111	−2	1111 1110
		−1	1111 1111
		0	禁止
		1	0000 0001
		2	0000 0010
		⋮	⋮
		127	0111 1111

10.2 H.263

ITU-T H.263改善了H.261在低码率情况下的视频质量。

H.263视频编码器提供了一个与其他信号（如H.223等）复用的独立数字视频比特流。视频解码器执行与编码器相反的过程。表10-9中列出了H.263有关YCbCr视频数据的主要规格。除了这些规格外，编解码器双方也可能协商一个定制的图像尺寸。图3-7中给出了4：2：0的YCbCr采样。

如下所示，H.263支持7种帧（或图像）类型，其中前两种是强制的（基准H.263）。

- Intra或I帧：没有采用参考帧进行预测的帧。
- Inter或P帧：基于前一个帧预测得到的帧。
- PB帧和改进的PB帧：一个代表两个帧的帧，且是基于前一个帧预测得到的帧。
- B帧：基于两个参考帧（前一帧和后一帧）预测得到的帧。
- EI帧：有一个同步帧的帧，同步帧大小与其相同或者稍小。
- EP帧：有两个参考帧（前一帧和同步帧）的帧。

479～481

表10-9 基准H.263的YCbCr参数

参　数	16CIF	4CIF	CIF	QCIF	SQCIF
有效分辨率（Y）	1408×1152	704×576	352×288	176×144	128×96
帧率	29.97Hz				
YCbCr采样结构	4：2：0				
YCbCr编码形式	归一化量化PCM，每采样点8位表示				

10.2.1 视频编码层

图10-8给出了一个典型的H.263编码器结构图，其基本的功能有预测、块变换和量化。

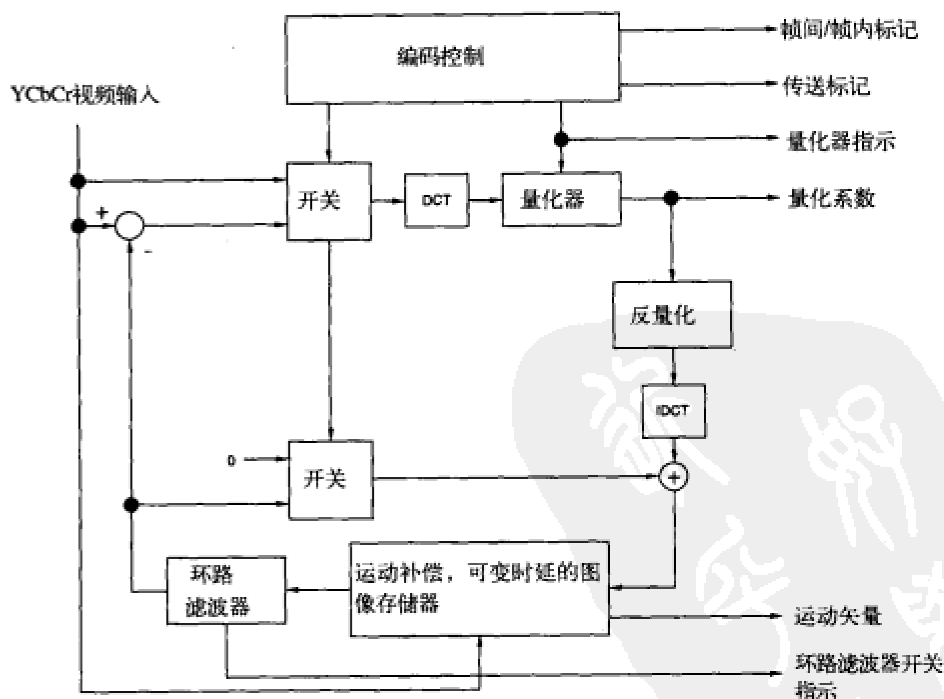


图10-8 典型的基准H.263编码器

预测误差或者输入图像被细分为多个8×8的块，这些块被分为传送的和非传送的。如图10-9所示，4个亮度块（Y）和两个空间上对应的色度块（Cb和Cr）组成了一个16×16的宏块。

H.263标准中没有给出模式选择和块传送的准则，这些准则可以作为编码策略的一部分动态地改变。编码器首先对发送块进行变换，然后对变换后的系数进行量化，最后对量化结果进行可变长编码。

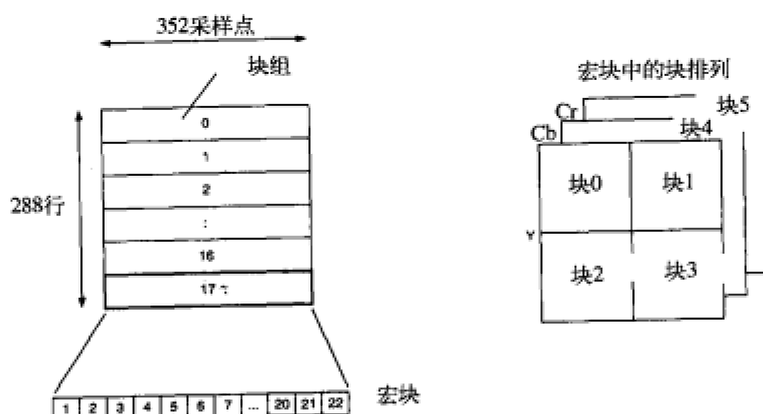


图10-9 H.263块组、宏块和块排列

1. 预测

预测发生在图像之间（inter-picture），可能包括运动补偿。采用预测的编码模式称为帧间编码（inter），而不采用预测的编码模式则称为帧内编码(intra)。

帧内编码模式用于图像层（对于intra是I帧，对于inter是P帧）或者P帧中的宏块层。在可选的PB帧模式中，B帧通常采用帧间编码模式。

2. 运动补偿

运动补偿在编码器中是可选的。解码器必须支持每个宏块拥有一个运动矢量（在可选的高级预测或者去块滤波器模式中，每个宏块有一个或4个运动矢量）。

在可选的PB帧模式中，每一个宏块可以有一个附加的矢量。在可选的改进的PB帧模式中，每一个宏块可以包含一个附加的前向运动矢量。在可选的B帧模式中，宏块在传送时可以有一个前向和后向运动矢量。

对于基准H.263，运动矢量是受限的，即所有被用作参考的采样点都必须在已编码的图像区域内，而对于许多可选模式，这一限制已经去除了。运动矢量的水平和垂直分量必须为−16~15.5范围内的整数值或半整数值。某些可选模式将这一范围增大到[−31.5，+31.5]或者[−31.5，+30.5]。

运动矢量的水平或垂直分量如果为正值，则表示预测是基于前一帧图像中位于被预测采样点右方或者下方的采样点。然而，对于B帧中的后向运动矢量，运动矢量的水平或垂直分量如果为正值，表示预测是基于下一帧图像中位于被预测的采样点左方或者上方的采样点。

3. 量化

对于第一个intra系数，量化器的数目是1，其他所有系数的量化器数目为31。在一个宏块中，除了intra-DC外，所有系数都采用相同的量化器。通常intra-DC采用的步长为8，且量化采用的是无死区的线性量化器。而其他系数则采用中心死区为0，步长为2，4，6，…，62之间的任意值的线性量化器量化。

4. 编码控制

H.263标准采用了多个参数来控制编码视频数据的产生速率，这些参数包括：编码器预处理过程、量化器、块重要性准则和时域亚采样等。但是，这些参数并没有作为其一部分而包含在H.263标准中。

482

5. 强制更新

这个功能是通过强制使用编码算法中的帧内模式（Intra）而实现的。为了防止IDCT变换的误

差累积或者其他一些差错的影响，源编码采取了一种强制更新的方法。它规定每个宏块至少每传送132次就要以帧内模式传送一次（更新一次），以防止编码中的误码累积。

10.2.2　视频比特流

除非特别声明，视频比特流的最高有效位MSB总是最先传送。MSB位于编码表的第1位，即最左边一位。同样，除非特别声明，所有的备用位均设置为1。

H.263的视频流复用结构划分为四个层次，从上到下依次是：

- ❑ 图像层P（Picture）
- ❑ 块组层GOB（Group of Blocks）或者条带（Slice）
- ❑ 宏块层MB（Macroblock）
- ❑ 块层B（Block）

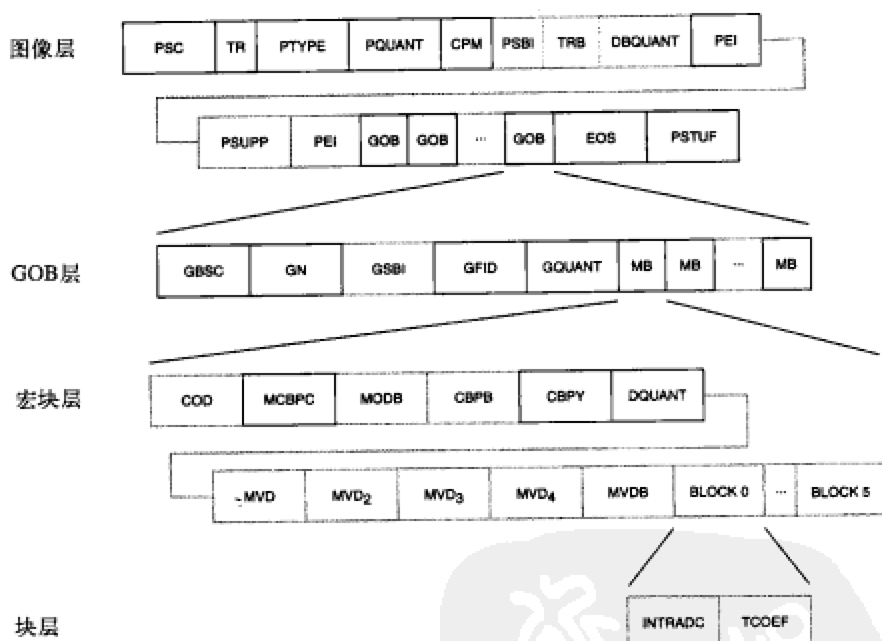


图10-10　基准H.263的视频比特流层次结构（在图像层无可选项PLUSPTYPE相关的字段）

1. 图像层

483～484

如图10-10所示，图像数据由图像头、块组数据和序列结束（EOS，end-of-sequence）标识以及填充位（PSTUF）组成。被丢弃图像的图像头不需要传送。

❑ **图像起始码（PSC）**

PSC（Picture Start Code）长度为22位，值为0000 0000 0000 0000 1 00000。PSC必须是字节对齐的，因此，起始码之前必须添加0~7个0位，以确保起始码的第一位是字节的第一位，即最高有效位。

❑ **时间参考（TR）**

TR（Temporal Reference）是一个8位的二进制数，表示256种可能值。TR的生成是通过将前一帧传送的图像中的TR值递增1，然后再加上非传送图像（在29.97Hz频率下）数目而得到的。运算仅在8个LSB上进行。

如果指定了定制图像时钟频率（PCF），那么采用扩展时间参考（ETR）和TR共同组成一个10位的编码，其中TR存储在8个LSB上，ETR存储在两个MSB上。在这种情况下，运算在10个LSB上进行。

485

在PB帧和改进的PB帧模式中，TR仅对P帧有效。

❑ 类型信息（PTYPE）

PTYPE（Type Information）采用13位来表示有关图像的信息：

位1表示1

位2表示0

位3表示分割屏幕指示器

0＝关，1＝开

位4表示文档相机指示器

0＝关，1＝开

位5表示冻结图像解除

0＝关，1＝开

位6~8表示源格式

000＝保留

001＝SQCIF

010＝QCIF

011＝CIF

100＝4CIF

101＝16CIF

110＝保留

111＝扩展的PTYPE

如果位6~8不是111，那么下面5位将出现在PTYPE中：

位9表示图像编码类型

0＝帧内，1＝帧间

位10表示可选的非限制运动矢量模式

0＝关，1＝开

位11表示可选的基于语法的算术编码模式

0＝关，1＝开

位12表示可选的高级预测模式

0＝关，1＝开

位13表示可选的PB帧模式

0＝正常图像，1＝PB帧

如果第9位设置位0，那么第13位必须设置为0。第10~13位是编码器和解码器之间协商的可选模式。

❑ 量化器信息（PQUANT）

PQUANT（Quantizer Information）是一个5位二进制编码（取值为1~31），它表示图像采用的量化器，PQUANT一直有效，直到被下一个GQUANT或者DQUANT所取代。

❑ 连续出现多点（CPM）

CPM（Continuous Presence Multipoint）是一个一位的值，表示编码器是否采用可选的连续出现多点和视频复用模式，0＝关，1＝开。如果PLUSPTYPE不存在，那么CPM紧接着PQUANT，否则CPM将紧接着PLUSPTYPE。

❑ 图像子比特流指示器（PSBI）

PSBI（Picture Sub-Bitstream Indicator）是一个可选的2位二进制编码，只有当CPM指示采用可选的连续存在多点和视频复用模式时，PSBI才会出现。

486

❑ PB帧中B帧时间参考（TRB）

当PTYPE或者PLUSTYPE指示帧类型为PB帧或者改进的PB帧时，PB帧中的B帧TRB才会出现。TRB是一个3位或者5位表示数[$n+1$]的二进制编码，其中n为从最近的I帧（或P帧）或者PB帧（或改进的PB帧）的P部分开始和PB帧（或改进的PB帧）的B部分之前非传送图像的数目（时钟频率为29.97Hz或者由CPCFC指示的定制图像时钟频率）。当采用定制图像时钟频率时，TRB的值扩展到5位。

当图像时钟频率为29.97Hz时，最大非传送图像数目为6，而采用定制图像时钟频率时，这个数为30。

❑ PB帧中B帧量化器信息（DBQUANT）

当PTYPE或者PLUSTYPE指示帧类型为PB帧或改进的PB帧时，将会出现DBQUANT（Quantizer Information for B Frames in PB Frames）。DBQUANT是一个2位码字，表示QUANT和BQUANT之间的关系，如表10-10所示。除法运算采用截切方式。BQUANT范围为1~31。如果运算结果低于1或者高于31，将BQUANT分别截切到1和31。

表10-10　基准H.263中DBQUANT编码和QUANT/BQUANT关系

DBQUANT	BQUANT	DBQUANT	BQUANT
00	(5*QUANT)/4	10	(7*QUANT)/4
01	(6*QUANT)/4	11	(8*QUANT)/4

❑ 额外插入信息（PEI）

PEI（Extra Insertion Information）采用1位表示，当设置为1时表示接下来将出现补充的增强信息PSUPP数据域。

❑ 补充的增强信息（PSUPP）

如果PEI设置为1，在PEI之后将紧随一个9位的额外插入信息，其中8位为PSUPP（Supplemental Enhancement Information），第9位为PEI位，表示在之后是否有另一个9位的额外插入信息，如此反复。

❑ 序列结束（EOS）

EOS（End of Sequence）是一个值为0000 0000 0000 0000 1 11111的22位字码。EOS必须为字节对齐，通过在编码之前插入0~7个零位，从而确保EOS编码的第一位为字节的第一位，并且是最高有效位。

❑ 填充（PSTUF）

PSTUF（Stuffing）是一个零位的变长字码。PSTUF的最后一位必须是字节的最后一位并且是最低有效位。

2. 块组层

如图10-9所示，每一个图像被细分成多个GOB。对于SQCIF、QCIF和CIF分辨率图像，一个GOB由16行组成，对于4CIF分辨率和16CIF分辨率的图像，一个GOB分别由32行和64行组成。因此，一个SQCIF、QCIF和CIF图像分别包含6（96/16）个、9（144/16）个和18（288/16）个GOB，每个GOB有一列宏块数据。而一个4CIF图像则有18（576/32）个GOB，每一个GOB有两列宏块数据。一个16CIF图像也有18（1152/64）个GOB，每一个GOB带有4行的宏块数据。GOB的编号方

487

式是从图像的顶部开始编号为0，然后垂直向下递增。

如图10-10所示，每一个GOB的数据由GOB头和宏块数据组成。宏块数据以递增的宏块编号顺序传送。对于每一个图像中编号为0的GOB，编码器不需要传送GOB头。解码器可以通知编码器仅传送非空的GOB头。

❑ 块组起始码（GBSC）

GBSC（Group of Blocks Start Code）长度为17位，值为0000 0000 0000 0000 1。GBSC必须是字节对齐的，因此，必须在起始码之前加入0~7个零位，以确保起始码的第一位是字节的第一位，并且是最高有效位。

❑ 组编号（GN）

GN（Group Number）是一个5位二进制数，表示GOB的编号。组编号1~17用于标准图像格式。组编号1~24用于定制图像格式。组编号16~29被复制到条带头中。组编号30用于子比特流结束EOSBS（end of sub-bitstream）指示符中，而组编号31则用于序列结束码EOS（end of sequence）中。

❑ GOB子比特流指示器（GSBI）

GSBI（GOB Sub-Bitstream Indicator）是一个2位的二进制编码，表示当前GOB到下一个图像或者下一个GOB起始码之间的子比特流数目。GSBI只有在采用连续出现多点和视频复用（CPM）模式时才会出现。

❑ GOB帧ID（GFID）

GFID用来表示GOB中的帧ID，长度为2位。对于一个给定的帧，在每一个GOB头（或者条带头）中都必须是同一个值。通常，如果当前图像的PTYPE与前一个图像头中的PTYPE是相同的，那么当前帧的GFID必须与前一个帧的GFID一致；否则，当前帧的GFID必须采用与前一个帧GFID不同的值。

❑ 量化器信息（GQUANT）

GQUANT（Quantizer Information）是一个5位二进制数，取值范围为1~31，表示块组中采用的量化器。GQUANT一致有效，直到被下一个GQUANT或者DQUANT所取代。

3. 宏块层

如图10-9所示，每一个GOB被细分成多个宏块。每一个宏块由16×16的Y块和对应的8×8的Cb和Cr块组成。宏块的编号方式是从左到右、从上往下递增的。宏块数据按照递增的宏块编号顺序传送。

宏块的数据由一个宏块头和块数据组成（参见图10-10）。

❑ 编码宏块指示（COD）

COD（Coded Macroblock Indication）采用一位来表示块是否已被编码。0表示已经编码，1表示没有编码，并且剩下的宏块层为空。只有当图像不是采用帧间编码模式时COD才会出现。

如果没有编码，解码器解码时将宏块当作一个整块运动矢量为零且没有系数数据的帧间块（inter-block）来处理。

❑ 宏块类型和色度编码块模式（MCBPC）

MCBPC（Macroblock Type and Coded Block Pattern for Chrominance）是一个变长的码字，表示宏块层类型和Cb与Cr的编码块模式。

表10-11和表10-12中列出了MCBPC的码字。填充位也有一个码字，但是在解码器解码时应该丢弃。在某些情况下，填充位不能出现在图像的第一个宏块中，这是为了避免与起始码混淆。表10-13和表10-14中列出了宏块类型（MB类型）。

488

表10-11　基准H.263中I帧的MCBPC可变长编码表

宏块类型	CBPC（Cb，Cr）	编　码		
3	0，0	1		
3	0，1	001		
3	1，0	010		
3	1，1	011		
4	0，0	0001		
4	0，1	0000	01	
4	1，0	0000	10	
4	1，1	0000	11	
填充		0000	0000	1

表10-12　基准H.263中P帧的MCBPC可变长编码表

宏块类型	CBPC（Cb，Cr）	编　码			
0	0，0	1			
0	0，1	0011			
0	1，0	0010			
0	1，1	0001	01		
1	0，0	011			
1	0，1	0000	111		
1	1，0	0000	110		
1	1，1	0000	0010	1	
2	0，0	010			
2	0，1	0000	101		
2	1，0	0000	100		
2	1，1	0000	0101		
3	0，0	0001	1		
3	0，1	0000	0100		
3	1，0	0000	0011		
3	1，1	0000	011		
4	0，0	0001	00		
4	0，1	0000	0010	0	
4	1，0	0000	0001	1	
4	1，1	0000	0001	0	
填充		0000	0000	1	
5	0，0	0000	0000	010	
5	0，1	0000	0000	0110	0
5	1，0	0000	0000	0111	0
5	1，1	0000	0000	0111	1

表10-13　基准H.263中宏块类型和标准帧包含的数据

帧类型	宏块类型	名　称	COD	MCBPC	CBPY	DQUANT	MVD	MVD_{2-4}
inter	没有编码	—	×					
inter	0	inter	×	×	×		×	
inter	1	inter+q	×	×	×	×	×	
inter	2	inter4v	×	×	×		×	×
inter	3	intra	×	×	×			

（续）

帧类型	宏块类型	名　称	COD	MCBPC	CBPY	DQUANT	MVD	$MVD_{2\text{-}4}$
inter	4	intra＋q	×	×	×	×		
inter	5	inter4v＋q	×	×	×	×	×	×
inter	填充	—	×	×				
intra	3	intra		×	×			
intra	4	intra＋q		×	×	×		
intra	填充	—		×				

表10-14a 基准H.263中宏块类型和PB帧包含的数据

帧类型	宏块类型	名　称	COD	MCBPC	MODB	CBPY
inter	没有编码	—	×			
inter	0	inter	×	×	×	×
inter	1	inter+q	×	×	×	×
inter	2	inter4v	×	×	×	×
inter	3	intra	×	×	×	×
inter	4	intra+q	×	×	×	×
inter	5	inter4v+q	×	×	×	×
inter	填充	—	×	×		

表10-14b 基准H.263中宏块类型和PB帧包含的数据

帧类型	宏块类型	名　称	CBPB	DQUANT	MVD	MVDB	$MVD_{2\text{-}4}$
inter	没有编码	—					
inter	0	inter	×		×	×	
inter	1	inter＋q	×	×	×	×	
inter	2	inter4v	×		×	×	×
inter	3	intra	×		×	×	
inter	4	intra＋q	×	×	×	×	
inter	5	inter4v＋q	×	×	×	×	×
inter	填充	—					

色度编码块模式（CBPC）表示对于Cb或者Cr是否有一个非intra-DC变换系数被传送。当CBPC为1时，表示在块中（Cb或者Cr）出现了一个非intra-DC系数。

❑ B块的宏块模式（MODB）

对于宏块类型为0~4的宏块，如果PTYPE指示为PB帧，将出现B块的宏块模式MODB（Macroblock Mode for Blocks）。MODB是一个变长的码字，表示对于当前宏块，B系数或运动矢量是否被传送。表10-15列出了MODB的码字。对于改进的PB帧，MODB采用了与PB帧不一样的编码方式。

表10-15 基准H.263的MODB可变长编码表

CBPB	MVDB	编　码
		0
	×	10
×	×	11

❑ B块编码块模式（CBPB）

如果MODB指示有CBPB，将出现一个6位长的CBPB（Coded Block Pattern for B Block）。CBPB指示宏块中哪一个块至少有一个变换系数被传送。模式编码方式如下：

$P_0P_1P_2P_3P_4P_5$

对于块[*n*]，若任意系数出现，则$P_n=1$，否则$P_n=0$。图10-9中给出了块编号方式（十进制）。

❑ 亮度编码块模式（CBPY）

CBPY（Coded Block Pattern for Luminance）是一个变长的码字，指定宏块中至少有一个非intra-DC变换系数被传送的Y块。然而，在先进帧内编码模式中，intra-DC采用与其他系数相同的方式来指示。

表10-16列出了CBPY的编码。对Y块，若任意非intra-DC系数出现，则Yn＝1。图10-9中给出了Y块的编号方式（十进制）。

表10-16 基准H.263中CBPY的可变长编码表

CBPY（Y0，Y1，Y2，Y3）		编 码	
intra	inter		
0，0，0，0	1，1，1，1	0011	
0，0，0，1	1，1，1，0	0010	1
0，0，1，0	1，1，0，1	0010	0
0，0，1，1	1，1，0，0	1001	
0，1，0，0	1，0，1，1	0001	1
0，1，0，1	1，0，1，0	0111	
0，1，1，0	1，0，0，1	0000	10
0，1，1，1	1，0，0，0	1011	
1，0，0，0	0，1，1，1	0001	0
1，0，0，1	0，1，1，0	0000	11
1，0，1，0	0，1，0，1	0101	
1，0，1，1	0，1，0，0	1010	
1，1，0，0	0，0，1，1	0100	
1，1，0，1	0，0，1，0	1000	
1，1，1，0	0，0，0，1	0110	
1，1，1，1	0，0，0，0	11	

❑ 量化器信息（DQUANT）

DQUANT是一个两位的码字，表示QUANT中值的改变。表10-17给出了该码字的不同值。

QUANT取值范围为1~31。如果DQUANT指示QUANT的值发生改变，并且改变后的结果小于1或者大于31，那么分别将其设置为1和31。

489

表10-17 基准H.263中QUANT差分值的DQUANT编码

QUANT的差分值	DQUANT	QUANT的差分值	DQUANT
−1	00	1	10
−2	01	2	11

❑ 运动矢量数据（MVD）

在PB帧模式中，所有的宏块间和块内都包含MVD（Motion Vector Data）。

运动矢量数据由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了变长码。对于非限制运动矢量模式，可能会采用其他运动矢量编码方式。

❑ 运动矢量数据

在高级预测模式或者去块滤波器模式过程中，如果PTYPE和MCBPC指示三个码字MVD_2、MVD_3和MVD_4存在，这三个运动矢量数据将会出现。每一个运动矢量数据都由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了这些变长码。

❑ B宏块运动矢量数据（MVDB）

在PB帧和改进的PB帧模式过程中，如果MODB指示MVDB存在，MVDB将会出现。MVDB由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了这些变长码。

表10-18 基准H.263中MVD、$MVD_{2\text{-}4}$和MVDB的可变长编码表

矢量差分		编码			
−16	16	0000	0000	0010	1
−15.5	16.5	0000	0000	0011	1
−15	17	0000	0000	0101	
−14.5	17.5	0000	0000	0111	
−14	18	0000	0000	1001	
−13.5	18.5	0000	0000	1011	
−13	19	0000	0000	1101	
−12.5	19.5	0000	0000	1111	
−12	20	0000	0001	001	
−11.5	20.5	0000	0001	011	
−11	21	0000	0001	101	
−10.5	21.5	0000	0001	111	
−10	22	0000	0010	001	
−9.5	22.5	0000	0010	011	
−9	23	0000	0010	101	
−8.5	23.5	0000	0010	111	
−8	24	0000	0011	001	
−7.5	24.5	0000	0011	011	
−7	25	0000	0011	101	
−6.5	25.5	0000	0011	111	
−6	26	0000	0100	001	
−5.5	26.5	0000	0100	011	
−5	27	0000	0100	11	
−4.5	27.5	0000	0101	01	
−4	28	0000	0101	11	
−3.5	28.5	0000	0111		
−3	29	0000	1001		
−2.5	29.5	0000	1011		
−2	30	0000	111		
−1.5	30.5	0001	1		
−1	31	0011			
−0.5	31.5	011			
0		1			

（续）

矢量差分		编　码			
0.5	−31.5	010			
1	−31	0010			
1.5	−30.5	0001	0		
2	−30	0000	110		
2.5	−29.5	0000	1010		
3	−29	0000	1000		
3.5	−28.5	0000	0110		
4	−28	0000	0101	10	
4.5	−27.5	0000	0101	00	
5	−27	0000	0100	10	
5.5	−26.5	0000	0100	010	
6	−26	0000	0100	000	
6.5	−25.5	0000	0011	110	
7	−25	0000	0011	100	
7.5	−24.5	0000	0011	010	
8	−24	0000	0011	000	
8.5	−23.5	0000	0010	110	
9	−23	0000	0010	100	
9.5	−22.5	0000	0010	010	
10	−22	0000	0010	000	
10.5	−21.5	0000	0001	110	
11	−21	0000	0001	100	
11.5	−20.5	0000	0001	010	
12	−20	0000	0001	000	
12.5	−19.5	0000	0000	1110	
13	−19	0000	0000	1010	
13.5	−18.5	0000	0000	1000	
14	−18	0000	0000	0110	
14.5	−17.5	0000	0000	0100	
15	−17	0000	0000	0100	
15.5	−16.5	0000	0000	0011	0

4. 块层

在非PB帧模式下，一个宏块由4个Y块和一个Cb块、一个Cr块组成（参见图10-9）。如图10-10所示，一个8样本×8行的块的数据由intra-DC系数和变换系数的码字组成。图10-9中给出了块传送的顺序。

而在PB模式下，一个宏块则由4个Y块和一个Cb块、一个Cr块和6个B块的数据组成。

量化的DCT系数将按图7-57中的顺序传送。在修改量化（modified quantization）模式中，量化的DCT系数将按图7-58中的顺序传送。

❑ 块内DC系数（intra-DC）

intra-DC（DC Coefficient for Intra-Blocks）是一个8位的码字。表10-19中列出了intra-DC的值及其对应的重构非零值。

在非PB帧模式下，如果MCBPC指示宏块的类型为3或者4，intra-DC系数将出现在宏块中的每一个块中。而在PB帧模式下，如果MCBPC指示宏块的类型为3或者4，intra-DC系数将出现在宏块中的每一个P块中（intra-DC系数不会出现在B块中）。

表10-19 基准H.263中intra DC的重构非零值

intra DC值	重构非零值
0000 0000	不采用
0000 0001	8
0000 0010	16
0000 0011	24
⋮	⋮
0111 1111	1016
1111 1111	1024
1000 0001	1032
⋮	⋮
1111 1101	2024
1111 1110	2032

❑ 变换系数（TCOEF）

在非PB帧模式下，TCOEF（Transform Coefficient）由MCBPC或者CBPY决定是否存在。在PB帧模式下，TCOEF由CBPB决定是否在B块中出现。

事件由最后一个非零系数指示（如果块中还有其他非零系数，LAST＝0；否则LAST＝1）、非零系数前连续零的个数（RUN）和非零系数值（LEVEL）组合而成。

如表10-20所示，大多数常见事件采用变长码编码。s位指示非零系数的符号，0表示正，1表示负。

490～494

表10-20 基准H.263中TCOEF的可变长编码表

LAST	RUN	|LEVEL|	编码		
0	0	1	10s		
0	0	2	1111	s	
0	0	3	0101	01s	
0	0	4	0010	111s	
0	0	5	0001	1111	s
0	0	6	0001	0010	1s
0	0	7	0001	0010	0s
0	0	8	0000	1000	01s
0	0	9	0000	1000	00s
0	0	10	0000	0000	111s
0	0	11	0000	0000	110s
0	0	12	0000	0100	000s
0	1	1	110s		
0	1	2	0101	00s	
0	1	3	0001	1110	s
0	1	4	0000	0011	11s
0	1	5	0000	0100	001s

（续）

LAST	RUN	|LEVEL|	编码			
0	1	6	0000	0101	0000	s
0	2	1	1110	s		
0	2	2	0001	1101	s	
0	2	3	0000	0011	10s	
0	2	4	0000	0101	0001	s
0	3	1	0110	1s		
0	3	2	0001	0001	1s	
0	3	3	0000	0011	01s	
0	4	1	0110	0s		
0	4	2	0001	0001	0s	
0	4	3	0000	0101	0010	s
0	5	1	0101	1s		
0	5	2	0000	0011	00s	
0	5	3	0000	0101	0011	s
0	6	1	0100	11s		
0	6	2	0000	0010	11s	
0	6	3	0000	0101	0100	s
0	7	1	0100	10s		
0	7	2	0000	0010	10s	
0	8	1	0100	01s		
0	8	2	0000	0010	01s	
0	9	1	0100	00s		
0	9	2	0000	0010	00s	
0	10	1	0010	110s		
0	10	2	0000	0101	0101	s
0	11	1	0010	101s		
0	12	1	0010	100s		
0	13	1	0001	1100	s	
0	14	1	0001	1011	s	
0	15	1	0001	0000	1s	
0	16	1	0001	0000	0s	
0	17	1	0000	1111	1s	
0	18	1	0000	1111	0s	
0	19	1	0000	1110	1s	
0	20	1	0000	1110	0s	
0	21	1	0000	1101	1s	
0	22	1	0000	1101	0s	
0	23	1	0000	0100	010s	
0	24	1	0000	0100	011s	
0	25	1	0000	0101	0110	s
0	26	1	0000	0101	0111	s
1	0	1	0111	s		
1	0	2	0000	1100	1s	
1	0	3	0000	0000	101s	

（续）

LAST	RUN	|LEVEL|	编	码		
1	1	1	0011	11s		
1	1	2	0000	0000	100s	
1	2	1	0011	10s		
1	3	1	0011	01s		
1	4	1	0011	00s		
1	5	1	0010	011s		
1	6	1	0010	010s		
1	7	1	0010	001s		
1	8	1	0010	00s		
1	9	1	0001	1010	s	
1	10	1	0001	1001	s	
1	11	1	0001	1000	s	
1	12	1	0001	0111	s	
1	13	1	0001	0110	s	
1	14	1	0001	0101	s	
1	15	1	0001	0100	s	
1	16	1	0001	0011	s	
1	17	1	0000	1100	0s	
1	18	1	0000	1011	1s	
1	19	1	0000	1011	0s	
1	20	1	0000	1010	1s	
1	21	1	0000	1010	0s	
1	22	1	0000	1001	1s	
1	23	1	0000	1001	0s	
1	24	1	0000	1000	1s	
1	25	1	0000	0001	11s	
1	26	1	0000	0001	10s	
1	27	1	0000	0001	01s	
1	28	1	0000	0001	00s	
1	29	1	0000	0100	100s	
1	30	1	0000	0100	101s	
1	31	1	0000	0100	110s	
1	32	1	0000	0100	111s	
1	33	1	0000	0101	1000	s
1	34	1	0000	0101	1001	s
1	35	1	0000	0101	1010	s
1	36	1	0000	0101	1011	s
1	37	1	0000	0101	1100	s
1	38	1	0000	0101	1101	s
1	39	1	0000	0101	1110	s
1	40	1	0000	0101	1111	s
ESC			0000	011		

（LAST，RUN，LEVEL）的其他组合采用一个22位的字来编码：其中7位表示转义（ESC），

1位为LAST，6位为RUN，剩余8位为LEVEL。表10-21中给出了RUN和LEVEL的编码。编码1000 0000被禁止使用，除非在修改的量化模式下。

表10-21　基准H.263的RUN、LEVEL编码

RUN	编　码	LEVEL	编　码
0	0000 00	−128	禁止
1	0000 01	−127	1000 0001
⋮	⋮	⋮	⋮
63	1111 11	−2	1111 1110
		−1	1111 1111
		0	禁止
		1	0000 0001
		2	0000 0010
		⋮	⋮
		127	0111 1111

除了intra-DC之外，其他所有系数都有一个重构非零值，取值范围为−2048~2047。重构非零值采用下面的方程恢复，其结果要进行削波处理。

若 LEVEL＝0，REC＝0

若 QUANT＝奇数：

$$|REC| = QUANT \times (2 \times |LEVEL| + 1)$$

若 QUANT＝偶数：

$$|REC| = QUANT \times (2 \times |LEVEL| + 1) - 1$$

在计算完|REC|之后，将添加其符号以获得REC值。Sign(LEVEL)将由表10-20中TCOEF编码中的s位指定。

$$REC = sign(LEVEL) \times |REC|$$

495~497

对于intra-DC块，重构非零值为：

$$REC = 8 \times LEVEL$$

5. PLUSPTYPE图像层选项

PLUSPTYPE由PTYPE的6~8位决定是否出现，用来确定是否采用H.263的版本2可选项。当PLUSPTYPE出现时，PLUSPTYPE及其相关的字段紧跟在PTYPE之后，并且在PQUANT之前。

如果PLUSPTYPE出现，CPM将紧跟在PLUSPTYPE之后；否则，CPM紧跟在PQUANT之后。PSBI总是紧跟在CPM之后（前提是CPM＝1）。

PLUSPTYPE是一个12位或者30位的码字，由三个子字段组成：UFEP、OPPTYPE和MPPTYPE。图10-11中给出了对PLUSPTYPE及其相关字段的说明。

❑ *更新全扩展的PTYPE*（UFEP）

UFEP是一个3位的码字，它由PTYPE中的“扩展 PTYPE”决定是否出现。

当UFEP值为000时，图像头中仅仅包含MPPTYPE。

当UFEP值为001时，图像头中同时包含OPPTYPE和MPPTYPE。如果图像类型为intra或者EI，UFEP字段必须设置为001。

此外，如果PLUSPTYPE出现在连续序列图像的每一帧中，那么每隔5秒钟或者每隔5帧（两者中较大的优先），UFEP字段必须设置为001。在易错环境中，UFEP字段应该更频繁地设置为001。

498~501

除000和001之外的其他值，目前暂时保留。

❑ PLUSPTYPE的可选部分（OPPTYPE）

OPPTYPE字段包含的特征不会一帧一帧地变化，也就是说这个字段在所有帧中都是一致的。如果UFEP设置为001，下面的18位将出现在OPPTYPE中。

位1~3为源格式

000＝保留

001＝SQCIF

010＝QCIF

011＝CIF

100＝4CIF

101＝16CIF

110＝定制源格式

111＝保留

位4为定制图像时钟频率

0＝标准，1＝定制

位5表示非限制运动矢量模式（UMV）

0＝关，1＝开

位6表示基于语法的算术编码模式（SAC）

0＝关，1＝开

位7表示高级预测模式（AP）

0＝关，1＝开

位8表示高级帧内编码模式（AIC）

0＝关，1＝开

位9表示去块滤波器模式（DF）

0＝关，1＝开

位10表示条结构化模式（SS）

0＝关，1＝开

位11表示参考图像选择模式（RPS）

0＝关，1＝开

位12表示独立段解码模式（ISD）

0＝关，1＝开

位13表示交替Inter-VLC模式（AIV）

0＝关，1＝开

位14表示改进的量化模式（MQ）

0＝关，1＝开

位15为1

位16为0

位17为0

位18为0

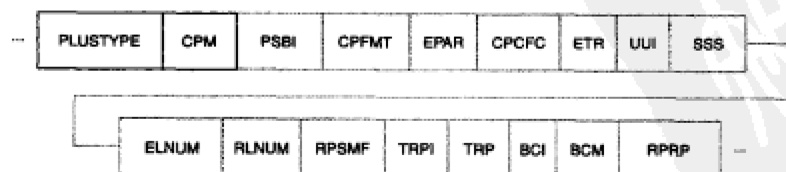


图10-11 H263的PLUSPTYPE及其相关字段

502

❑ PLUSPTYPE的必选部分（MPPTYPE）

无论UFEP的值为多少，下列9位总是出现在MPPTYPE中：

位1~3表示图像编码类型

000＝I帧（intra）

001＝P帧（inter）

010＝改进的PB帧

011＝B帧

100＝EI帧

101＝EP帧

110＝保留

111＝保留

位4表示参考图像重采样模式

0＝关，1＝开

位5表示降低分辨率更新模式（RPR）

0＝关，1＝开

位6表示取整类型模式（RTYPE）

0＝关，1＝开

位7为0

位8为0

位9为1

❑ 定制图像格式（CPFMT）

CPFMT（Custom Picture Format）是一个23位的码字。如果UFEP设置为001，并且PLUSPTYPE的OPPTYPE中前3位指定的源格式为定制图像格式，CPFMT将出现。

位1~4 为像素宽高比（PAR）编码

0000＝保留

0001＝1：1

0010＝12：11

0011＝10：11

0100＝16：11

0101＝40：33

0110~1100＝保留

1111＝扩展的像素宽高比EPAR

位5~13为图像宽度指示（PWI）

每行采样点数＝（PWI＋1）×4

位14为1

位15~23表示图像高度指示（PHI）

每帧扫描线数＝（PHI＋1）×4

❑ 扩展的像素宽高比（EPAR）

EPAR（Extended Pixel Aspect Ratio）是一个16位的值。如果CPFMT出现，且CPFMT的前4位指示采用扩展的像素宽高比，EPAR将出现。

位1~8表示PAR宽度

位9~16表示PAR高度

❑ 定制图像时钟频率编码（CPCFC）

CPCFC（Custom Picture Clock Frequency Code）是一个8位的值。只有当PLUSPTYPE出现、

UFEP设置为001且在PLUSPTYPE的OPPTYPE中第4位指示采用定制图像时钟频率时，CPCFC才会出现。定制的图像时钟频率（单位为赫兹）为：

1 800 000/（时钟因子×时钟转换因子）

位1 表示时钟转换因子编码

0＝1000，1＝1001

位2~8表示时钟因子

❑ *扩展时间参考*（ETR）

ETR（Extended Temporal Reference）是一个2位的值。当采用定制图像时钟频率时，ETR将会出现。它是10位TR值中的2个最高有效位。

503

❑ *无约束非限制运动矢量指示器*（UUI）

UUI（Unlimited Unrestricted Motion Vectors Indicator）是一个1位或者2位的可变长值，它指示运动矢量变化的有效范围。如果UFEP设置为001，且PLUSPTYPE的OPPTYPE中第5位指示采用非限制运动矢量模式，UUI将会出现。

UUI值为0时，表示运动矢量变化范围将根据表10-22和表10-23确定，而UUI值为01时，表示运动矢量在图像大小内是无约束的，即可以为图像大小内任意值。

表10-22 可选的水平运动矢量范围

图像宽度	水平运动矢量范围
4~352	−32，＋31.5
356~704	−64，＋63.5
708~1408	−128，＋127.5
1412~2048	−256，＋255.5

表10-23 可选的垂直运动矢量范围

图像高度	垂直运动矢量范围
4~288	−32，＋31.5
292~576	−64，＋63.5
580~1152	−128，＋127.5

❑ *条带结构化子模式位*（SSS）

SSS（Slice Structured Submode Bit）是一个2位的值。只有当UFEP设置为001，且在PLUSPTYPE的OPPTYPE中第10位指示采用条带结构化模式时，SSS才会出现。如果当前采用的是条带结构化模式，而UFEP不是001，那么最近的SSS值仍然保持有效。

位1表示矩形条带：0＝不是，1＝是。

位2表示任意条带顺序：0＝顺序的，1＝随机的。

❑ *增强层编号*（ELNUM）

ELNUM（Enhancement Layer Number）是一个4位的值，它只有在时域、信噪比（SNR）和空域可分级模式中出现。ELNUM标识一个特定的增强层。基本层之上的第一个增强层指派为增强层编号为2，基本层编号为1。

❑ *参考层编号*（RLNUM）

RLNUM（Reference Layer Number）是一个4位的值，它只有在时域、信噪比（SNR）和空域可分级模式下，且UFEP设置为001时出现。用作参考锚点（reference anchor）的帧的层编号由RLNUM确定。

❑ *参考图像选择模式标志*（RPSMF）

RPSMF（Reference Picture Selection Mode Flag）是一个3位的码字，它只有在参考图像选择模式下且UFEP设置为001时才会出现。当RPSMF出现时，它表示解码器需要哪一个反向信道消息。如果当前采用的是参考图像选择模式，而RPSMF没有出现，那么RPSMF最近的值仍然保持有效。

000~011＝保留

100＝ACK和NACK都不需要

101＝需要ACK
110＝需要NACK
504
111＝ACK和NACK都需要

❑ 预测时间参考指示（TRPI）

TRPI（Temporal Reference for Prediction Indication）是一个1位的值，它只有在参考图像选择模式中出现。当TRPI出现时，它表示接下来的TRP字段是否会出现。0＝TRP，字段不会出现，1＝TRP，字段出现。当图像头指示当前帧为I帧或者EI帧时，TRPI必须设置为0。

❑ 预测时间参考（TRP）

TRP（Temporal Reference for Prediction）是一个10位的值，对于除B帧之外的其他所有帧，TRP表示用于编码预测的时间参考。对于B帧，由TRP确定的时间参考将仅用于正向方向的编码预测。

如果没有采用定制时钟频率，那么参考图像的图像头中TRP的两位最高有效位将为0，而最低有效位包含8位TR值。如果采用定制时钟频率，TRP则为一个10位的值，由来自参考图像头中的ETR和TR串联组成。

如果TRP没有出现，前一个锚点图像将被用于预测，这与不是在参考图像选择模式中是一致的。TRP值将直到出现下一个PSC、GSC或者SSC时才有效。

❑ 反向信道消息指示（BCI）

BCI（Back-Channel Message Indication）是一个1位或者2位的可变长码字，它只有在可选的参考图像选择模式中出现。BCI值为1时表示有可选的反向信道消息（BCM）字段。而BCI值为01时，表示没有BCM字段或者BCM字段结束。当BCI出现时，BCM和BCI可以重复。

❑ 反向信道消息（BCM）

如果前面的BCI字段设置为1，可变长的BCM（Back-Channel Message）将会出现。

❑ 参考图像重采样参数（RPRP）

RPRP（Reference Picture Resampling Parameters）是一个可变长的字段，它只在可选的参考图像选择模式中出现。RPRP包含了参考图像重采样模式的相关参数。

10.2.3 可选的H.263模式

1. 非限制运动矢量模式

通常运动矢量的范围被限制在参考帧内，而在非限制运动矢量模式中取消了这种限制，运动矢量可以指向图像之外。当某个运动矢量所指的参考像素位于图像之外时，可以用边缘像素代替这个"不存在的像素"。这个边界像素的确定方法是：将运动矢量限制在编码图像区域内的最后一个全像素位置。对运动矢量的限制是在像素的基础上实现的，并且是分别针对运动矢量的水平分量和垂直分量进行的。

此外，非限制运动矢量模式扩展了运动矢量范围，因此，在该模式下可以采用更大的运动矢量（表10-22和表10-23）。对于分辨率较高的图像格式（如4CIF或者16CIF），这些大范围的运动矢量将很好地改善编码效率。对于其他图像格式，如果在图像边缘处存在运动（摄像机运动或者背景运动），也将获得明显的图像质量改善。

当在H.263版本2中采用非限制运动矢量模式时，将采用一种新的可逆变长编码（RVLC）用于运动矢量的编码，如表10-24所示。这些编码是单值的（single-valued），与此相对的是基线双值VLC。由于受到其可扩展性及其高实现代价的限制，双值编码不是很普遍，而RVLC实现相对简单。

表10-24中的每一行表示一个运动矢量差分（单位为半像素）。"...X1X0"表示在运动矢量差分绝对值的二进制表示中第一位1之后的所有位。s位表示运动矢量差分的符号：0表示为正，1表示为负。运动矢量差分的二进制表示采用比特交织，并且指示编码是否继续或者结束。最后一位如果为0，表示编码结束。

505

RLVC也可以增强对信道传输误码的错误恢复能力。解码可以通过在前向和反向方向处理运动矢量来执行。如果在某个方向上解码时探测到一个误码，解码器可以在反方向上继续进行下去，从而改善比特流的错误恢复能力。此外，运动矢量范围被扩展到[−256, +256]之间，具体值取决于图像大小。

表10-24　H.263运动矢量的可逆变长编码

半像素精度运动矢量差分的绝对值	编　码
0	1
1	$0s_00$
"x_0" +2(2−3)	$0x_01s0$
"x_1x_0" +4(4−7)	$0x_11x_01s0$
"$x_2x_1x_0$" +8(8−15)	$0x_21x_11x_01s0$
"$x_3x_2x_1x_0$" +16(16−31)	$0x_31x_21x_11x_01s0$
"$x_4x_3x_2x_1x_0$" +32(32−63)	$0x_41x_31x_21x_11x_01s0$
"$x_5x_4x_3x_2x_1x_0$" +64(64−127)	$0x_51x_41x_31x_21x_11x_01s0$
"$x_6x_5x_4x_3x_2x_1x_0$" +128(128−255)	$0x_61x_51x_41x_31x_21x_11x_01s0$
"$x_7x_6x_5x_4x_3x_2x_1x_0$" +256(256−511)	$0x_71x_61x_51x_41x_31x_21x_11x_01s0$
"$x_8x_7x_6x_5x_4x_3x_2x_1x_0$" +512(512−1023)	$0x_81x_71x_61x_51x_41x_31x_21x_11x_01s0$
"$x_9x_8x_7x_6x_5x_4x_3x_2x_1x_0$" +1024(1024−2047)	$0x_91x_81x_71x_61x_51x_41x_31x_21x_11x_01s0$
"$x_{10}x_9x_8x_7x_6x_5x_4x_3x_2x_1x_0$" +2048(2048−4095)	$0x_{10}1x_91x_81x_71x_61x_51x_41x_31x_21x_11x_01s0$

2. 基于语法的算术编码模式

在这个可选模式中，可变长编码被算术编码取代了。信噪比SNR和重构图像将会是一样的，但是比特率将会减少近5%，因为没有了每一个符号必须用固定长度比特数编码的需求。

在基于语法的算术编码模式中，图像层、块组层和宏块层的语法完全保持一致，块层的语法稍微改变了一点，即TCOEF实体数目可能会出现任意值。

值得注意的是，基于语法的算术编码模式并没有被普遍地采用。

506

3. 高级预测模式

通常16×16的宏块只有一个运动矢量，而在这种可选的高级预测模式下，每个宏块的4个Y块各有一个运动矢量。此外，在此模式下，对P帧中的Y块进行了交叠块运动补偿（overlapped block motion compensation，OBMC）处理。

如果对某一特定宏块只传送了一个运动矢量，那么这与传送四个值完全相同的矢量是等价的。如果四个运动矢量被用于一个宏块，那么第一个运动矢量将是MVD码字，并且应用到图10-9中的Y_1；第二个运动矢量为MVD_2码字，应用到Y_2；第三个运动矢量为MVD_3码字，应用到Y_3；第四个运动矢量为MVD_4码字，应用到Y_4。宏块中Cb和Cr的运动矢量将从这四个Y块运动矢量导出。

在高级预测模式下，编码器必须决定采用哪种类型的运动矢量。四个运动矢量将采用更多的位，但是它们将提供更好的预测。这种模式将改善图像间预测，从而通过减少方块效应（blocking artifacts），在相同的比特率下将获得明显的图像质量改善。

4. PB帧模式

与MPEG一样，H.263选择性地支持PB帧。如图10-12所示，在H.263中一个PB帧单元由两帧

组成，其中的P帧是由前一个P帧预测所得的，而B帧则是由前一个P帧和本PB帧中的P帧通过双向预测所得的结果。

在该编码选项下，在不大幅提高比特率的情况下图像速率获得了明显的提高。然而，在Annex M[①]中，提供了对改进的PB帧模式的支持。因此，仅仅为了保持与在采用Annex M之前设计的系统兼容，原始PB帧模式得以保留。

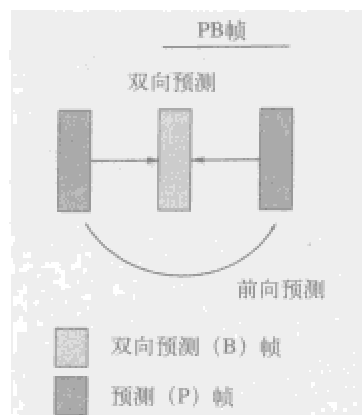


图10-12 基准H.263 PB帧

5. 连续出现多点和视频复用模式

在连续出现多点和视频复用模式下，一个单独的比特流最多可以复用4个独立的H.263子比特流。除非通过其他方法设置了不同的优先级约定，通常情况下，ID号最低的子比特流（通过SBI字段发送）被赋予了最高优先级。

上述特征被设计用于连续出现多点应用或分离逻辑信道无法获得的情形，这种情况下必须采用复用视频比特流。在H.324中将不采用连续出现多点和视频复用模式。

6. 前向纠错模式

前向纠错模式为传输H.263视频数据提供了前向纠错功能（编码和成帧）。在H.324中不采用该模式。

在前向纠错模式下，H.263的成帧和前向纠错编码与H.261中是完全一致的。

7. 高级帧内编码模式

高级帧内编码模式改善了宏块内的压缩效率。它采用来自相邻块内的块内预测、修改的intra-DCT系数反量化和单独的intra系数的VLC表。这种模式明显地改善了基准H.263帧内编码的压缩性能。

在高级帧内编码模式下，在宏块层中的MCBPC字段之后增加了一个附加的1位或者2位的变长码字（INTRA_MODE），用来指示预测模式：

0＝仅DC

10＝垂直DC和AC

507

11＝水平DC和AC

对于帧内编码块，如果预测模式仅为DC，采用图7-57中的锯齿形扫描顺序。如果预测模式为垂直DC和AC，采用图7-58中交替垂直扫描顺序。如果预测模式为水平DC和AC，采用图7-59中的交替水平扫描顺序。

① 一种ADSL技术。——译者注

对于非帧内编码块，将采用图7-57中的锯齿形扫描顺序。

8. 去块滤波器模式

在该可选模式中，编码环路中引入了一个去块滤波器。这个去块滤波器被应用到8×8块的边界处，用来减少方块效应。

去块滤波器的系数取决于宏块量化器的步长，大的系数被用于粗粒度的量化器。与高级预测模式一样，去块滤波器模式也允许每个宏块采用四个运动矢量，并且与非限制运动矢量模式一样，也允许运动矢量指向图像区域之外。由于没有采用高级预测模式中高计算代价的交叠运动补偿运算，因此保持了该模式最小的复杂性。

采用去块滤波器模式获得的结果是更好的预测和方块效应减少。

9. 条带结构化模式

在该可选模式中，条带层取代了GOB层。这种模式提供了错误恢复能力，使得比特流容易被分组传输传递机制所使用，并且最小化视频时延。

条带层由一个条带头和连续的完整宏块组成。在条带结构化模式中，有两个附加的模式可以用来反映传输的顺序（顺序的或者随机的）和条带的形状（矩形的或者非矩形的）。这些附加的模式使得条带结构更加灵活，从而能够被用于不同的应用。

10. 补充的增强信息

在该可选模式下，比特流中可能会包含额外的补充信息，以获得增强的显示能力。

典型的增强信息可以为视频复合提供全部或者部分图像冻结[①]、图像冻结释放或者色度抠像的能力。 508

在某些情况下，即便解码器无法使用这些补充的增强信息，它们仍然有可能会出现在比特流中。对此，解码器只是简单地丢弃这些补充信息，除非编解码器已经通过另外的方法协商以支持这种能力的需求。

11. 改进的PB帧模式

改进的PB帧模式是相对基准H.263的PB帧可选项的一种改进。这种模式允许在PB帧中对B帧进行前向、后向和双向预测。表10-25中给出了MODB字段变化的操作。

表10-25 改进的PB帧模式中MODB的H.263可变长编码表

CBPB	MVDB	编码	编码模式
		0	双向预测
x		10	双向预测
	x	110	前向预测
x	x	1110	前向预测
		11110	后向预测
x		11111	后向预测

改进的PB帧模式中的双向预测方法与PB帧模式的双向预测方法是一样的，唯一不同之处就是在改进的PB帧模式中不发送δ矢量。

在前向预测中，B宏块由前一个的P宏块预测得到，并且发送一个单独的运动矢量。

在后向预测中，预测的宏块等同于后一个P宏块，因此没有运动矢量被发送。

改进的PB帧不易受到可能发生在帧之间的变化的影响，如前一个P帧和当前PB帧之间的场景切变。

① 图像冻结为在视频播放过程中整帧或者部分图像冻结，其他部分继续播放的效果。——译者注

12. 参考图像选择模式

在基准H.263中，某一帧可能由前一帧预测得到。如果参考帧的某一部分由于信道误码或者分组丢失而被丢失了，由这些参考帧推导出来的帧的质量将会下降。采用参考图像选择模式，编码器将有可能选择采用哪一个参考帧来进行预测，从而最小化误码传播。

编码器和解码器利用四种反向信道消息信号（NEITHER、ACK、NACK和ACK＋NACK）来确定哪一个图像段将被用于预测。例如，解码器发送一个NACK消息给编码器，表示给定帧质量由于信道误码而下降了。因此，编码器可以选择不使用这一帧用于接下来的预测，转而采用另外的、质量没受影响的参考帧。这可以降低误码传播，从而在易错环境中保持较好的图像质量。

13. 时域、SNR和空域可分级模式

在这个可选模式中，分别支持三种不同的可分级模式：时域、SNR和空域。可分级使得解码器可以在多个质量层次上解码视频序列。可分级是通过采用将图像层次化且将增强图像分成一个或多个层而实现的。最低层被称为基本层。

基本层是一个可单独解码的比特流。增强层可以与基本层一起解码，从而增加图像速率，提高图像质量或者增大图像尺寸。

时域可分级是通过双向预测图像（或B帧）来达到的。在这种模式下，预测可以是基于基本层中的前一个图像或者后一个图像，或者两者都采用。其结果是获得比P帧更好的压缩比。这些B帧不同于PB帧或者改进的PB帧中的B图像部分，主要在于它们在比特流中是独立的实体。

SNR可分级依靠增强信息来提高图像质量而不增加分辨率。由于压缩引入了块效应，因此，

509

解码后图像与原始图像之间不可避免地存在差异，这种差异称为编码误差。正常情况下，编码误差在编码器中被丢弃，并且永远不会恢复。而在SNR可分级模式下，编码误差被发送到解码器，使得解码图像获得增强。这些额外的数据主要用于提高图像的SNR，这也就是术语"SNR可分级"的由来。

空域可分级与SNR可分级非常相关。唯一的不同就是在参考层中图像被用于预测空域增强层中图像之前，它被插入了1/2的垂直或者水平运动矢量（一维空域可分级），或者两者都有（二维空域可分级）。除了上采样过程之外，空域可分级图像的处理和语法与SNR可分级图像是一样的。

由于在采用SNR可分级的帧与采用空域可分级的帧之间有非常小的句法差异，因此，采用这两种可分级模式的帧分别被称为EI帧和EP帧。

在基本层中，用于在EI帧或者EP帧中向上预测的帧可能是I帧、P帧、PB帧中的P部分或者改进的PB帧的中P部分（但是不可能是B帧，PB帧中的B部分或者改进的PB帧中的B部分）。

该模式对于带宽容量变化的网络非常有用。

14. 参考图像重采样模式

在该可选模式中，在将参考帧用于预测之前，需要对参考帧进行重采样，以获得不同的图像大小。

参考图像重采样模式使得编码器有一个与被预测的帧不同的源参考格式。它也可以通过改变（warping）参考帧的形状、大小和位置来进行全局运动估计，或者旋转矢量估计。

15. 降低分辨率更新模式

该可选模式允许编码器发送某一帧以较低分辨率编码的更新信息，而对于参考帧，则仍然保持较高的分辨率，从而获得一个最终分辨率较高的帧。

降低分辨率更新模式最适合用于编码高动态场景，对于场景中运动部分，允许编码器增加其

510

帧率，而对于场景中的静态区域则维持较高的分辨率。

降低分辨率更新模式的语法与基准H.263是一致的，但是语义的解释有所差别。在该模式下，

宏块的维度翻倍了，所以宏块数据大小将是不采用这种模式时的四分之一。因此，在所有维度上的运动矢量也必须翻倍。为了产生最终的图像，宏块将上采样到扩展的分辨率。在向上采样以后，将全分辨率帧与运动补偿帧相加，以生成全分辨率帧作为后续帧的参考。

16. 独立段解码模式

在该可选模式中，将图像的段边界作为图像的边界进行处理，即允许在段边界处无数据依赖。

采用这种模式可以防止差错的传播，提供错误恢复和恢复的能力。这种模式最适合用于条带层，例如，可以调整条带大小以匹配特定的分组大小。

17. 交替Inter-VLC模式

当采用交替Inter-VLC模式时，用于高级帧内编码模式中的intra-VLC表同样可以用于块间编码。

当采用小的量化器步长时，通常出现在块内部的大量化系数和小游程越来越频繁地出现在块之间。当可以比特节省，并且解码器可以探测到采用内部量化DCT系数表时，解码器将采用intra表。解码器将首先试图采用inter表解码量化的系数。如果导致寻址系数在8×8块的64个系数之外，解码器将采用intra表。

18. 修改的量化模式

该可选模式将改善对编码的比特率控制，减少CbCr量化误差，扩展DCT系数的范围以及对系数值给予一定的限制。

在基准H.263中，可以在宏块层修改量化器值。然而，仅仅允许对最近的量化器的值进行小的调整（±1或±2）。修改的量化模式允许将量化器的值修改为任意值。

在基准H.263中，Y块和CbCr块量化器是一致的。在修改的量化模式中，通过对相对于Y块的Cb和Cr块采用一个较小的量化器步长，同样提高了Cb和Cr块的图像质量。

在基准H.263中，当采用一个小于8位的量化器时，超出[-127, +127]范围的量化系数将被削波。而在修改的量化模式下，量化系数被允许超出[-127, +127]范围。因此，当选择一个非常精细的量化步长时，Y块的质量将获得较大的提升。

19. 增强的参考图像选择模式

增强的参考图像选择（ERPS）模式提供了增强的编码效率和错误恢复能力。它管理着一个存储图像的多图像缓存器。

20. 数据分区的条带模式

数据分区的条带（DPS）模式提供了增强的错误恢复能力。它将头数据和运动矢量数据从DCT系数数据中分离出来，并且通过采用可逆的表示方法保护运动矢量数据。 511

21. 额外的补充增强信息说明

额外的补充增强信息说明提供了向后兼容的增强，如下所示。

（1）采用一个特定固定点IDCT的指示。

（2）图像消息，包括的消息类型有：任意二进制数据，文本（任意、版权、标题、视频描述或URL），图像头副本（有可靠时间参考的当前、前一个、下一个图像或者无可靠时间参考的下一个图像），隔行扫描字段（顶部或者底部），备用参考图像标识。

10.2.4 档次

档次定义了与特定应用相关的常用配置语法（如算法）。

1. 档次0

基本档次或者档次0，不采用任何可选模式。

2. 档次1

档次1（H.320编码效率版本2向后兼容档次）提供与H.242和H.320的兼容性。它由档次0与下面的模式共同组成：高级帧内编码，去块滤波器，补充的增强信息——全图像冻结，修改的量化。

3. 档次2

档次2（版本1向后兼容档次）为H.263的第一个版本提供增强的编码效率。它由档次0加上下面的模式组成：

512

- 高级预测

4. 档次3

档次3（版本2交互式的和流化无线档次）为无线设备提供增强的编码效率性能和增强的错误恢复能力。它由档次0加上下列模式组成：高级帧内编码，去块滤波器，条带结构化，修改的量化。

5. 档次4

档次4（版本3交互式的和流化无线档次）为无线设备提供可增强的编码效率性能和增强的错误恢复能力。它由档次0和档次3加上下列模式组成：数据分区条带，补充的增强信息——前一个图像头副本。

6. 档次5

档次5（会话式高压缩档次）在不增加与采用B图像相关的时延和增加错误恢复特性的情况下，提供增强的编码效率。它由档次0、档次1和档次2加上下列模式组成：非限制运动矢量——UUI＝1，增强的参考图像选择。

7. 档次6

档次6（会话式Internet档次）在不增加与采用B图像相关的时延，增加某些适合于在IP网络上使用的错误恢复能力的情况下，提供增强的编码效率性能。它由档次0和档次5加上下面的模式组成：随机条带顺序的条带结构化，补充的增强信息——前一个图像头副本。

8. 档次7

档次7（会话式隔行扫描档次）为低时延应用提供增强的编码效率性能，并且支持隔行扫描视频源。它由档次0和档次5加上下面的模式组成：补充的增强信息——隔行扫描字段指示（对于240×288图像）。

9. 档次8

档次8（高延迟档次）为没有紧要时延约束的应用提供增强的编码效率性能。它由档次0和档次6加上下列模式组成：参考图像重采样，时域可分级——B图像。

10.2.5　级别

级别定义了一个档次中的各种参数（分辨率、帧率和比特率等）。

1. 级别10

最大支持176×144分辨率，最大比特率为64kbit/s。

2. 级别20

最大支持352×288分辨率，最大比特率为128kbit/s。

3. 级别30

513

最大支持352×288分辨率，最大比特率为384kbit/s。

4. 级别40

最大支持352×288分辨率，最大比特率为2Mbit/s。

5. 级别45

最大支持176×144分辨率，最大比特率为128kbit/s。

6. 级别50

最大支持352×288分辨率，最大比特率为4Mbit/s。

7. 级别60

最大支持720×288分辨率，最大比特率为8Mbit/s。

8. 级别70

最大支持720×576分辨率，最大比特率为16Mbit/s。

参考文献

1. *Efficient Motion Vector Estimation and Coding for H.263-Based Very Low Bit-Rate Video Compression*, by Guy Cote, Michael Gallant, and Faouzi Kossentini, Department of Electrical and Computer Engineering, University of British Columbia.
2. *H.263+: Video Coding at Low Bit-Rates*, by Guy Cote, Berna Erol, Michael Gallants, and Faouzi Kossentini, Department of Electrical and Computer Engineering, University of British Columbia.
3. ITU-T H.261, *Video Codec for Audiovisual Services at p ×64 kbits*, 3/93.
4. ITU-T H.263,*Video Coding for Low Bit-Rate Communication*, 01/2005

514



第11章　消费DV

DV（Digital Video，数字视频）格式由基于磁带的数字便携式摄像机使用，它基于IEC 61834（25Mbit/s比特率）和新近的SMPTE 314M和370M规范（25、50或者100Mbit/s比特率）。尽管消费DV与MPEG的I帧有许多地方相似，但是它所使用的压缩算法既不是M-JPEG也不是MPEG。消费DV使用一个可以编辑的专有压缩算法，因为这种算法是一种帧内技术。

数字化的视频在压缩之前存储在内存中。然后测量缓冲器中存储的两个场之间的相关性。如果相关性低，表示场间存在运动，两个场单独进行压缩。通常情况下，压缩的是整个帧。不管是哪一种情况，都采用基于DCT的压缩。

为了实现恒定的25、50或者100Mbit/s的比特率，DV使用自适应量化，即对每个帧使用合适的DCT量化表。

图11-1描述了磁带上所写的一个磁道的内容。ITI区（Insert and Track Information sector，插入和磁道信息区）包含磁道状态的信息，作为视频编辑期间的约定控制磁道。



图11-1　480i系统磁道分区的安排。每个磁道的总位数为134 975（134 850）bit，包括重写裕度。每个视频帧有10（12）个磁道。576i系统参数（如果不同）示范在括号中

音频区（audio sector）包含音频数据和辅助音频数据（auxiliary audio data，AAUX），如图11-2所示。

视频区（video sector）包含视频数据和辅助视频数据（auxiliary video data，VAUX），如图11-3所示。VAUX数据包括录制日期和时间、透镜孔径、快门速度、色平衡和其他摄像机设置。

子码区（subcode sector）存储一系列的信息，包括时间码、图文电视、多语言中的隐藏字幕、多语言中的对白字幕和卡拉OK歌词、标题、目录、章节等。子码区、AAUX数据和VAUX数据使用5字节的数据块，这样的数据块称为包组。

11.1　音频

音频帧以第1行开头（480i系统）或者第623行中间（576i系统）前面50个样本中的一个音频样本为起始点。

同步块号					
0	同步 2B	ID 3B	AAUX数据5B	音频数据 72B	内部奇偶校验 8B
1	同步 2B	ID 3B	AAUX数据5B	音频数据 72B	内部奇偶校验 8B
⋮					
8	同步 2B	ID 3B	AAUX数据5B	音频数据 72B	内部奇偶校验 8B
9	同步 2B	ID 3B	外部奇偶校验		内部奇偶校验 8B
⋮					
13	同步 2B	ID 3B	外部奇偶校验		内部奇偶校验 8B

图11-2 音频区同步块的结构

同步块号				
0	同步 2B	ID 3B	VAUX数据 77B	内部奇偶校验 8B
1	同步 2B	ID 3B	VAUX数据 77B	内部奇偶校验 8B
2	同步 2B	ID 3B	视频数据 77B	内部奇偶校验 8B
⋮				
136	同步 2B	ID 3B	视频数据 77B	内部奇偶校验 8B
137	同步 2B	ID 3B	VAUX数据 77B	内部奇偶校验 8B
138	同步 2B	ID 3B	外部奇偶校验	内部奇偶校验 8B
⋮				
148	同步 2B	ID 3B	外部奇偶校验	内部奇偶校验 8B

图11-3 视频区同步块的结构

每个磁道包含9个音频同步块，每个音频同步块包含5B的音频辅助数据（AAUX）和72B的音频数据，如图11-2所示。音频样本散乱地位于帧的各个磁道和数据同步块中。剩余的5个音频同步块用于错误校正。

2个44.1kHz、16位的通道要求大约1.64Mbit/s的数据传输率。4个32kHz、12位的通道要求大约1.536Mbit/s的数据传输率。2个48kHz、16位通道要求大约1.536Mbit/s的数据传输率。

11.1.1 IEC 61834

IEC 61834支持多种音频采样率：48kHz（16位，2通道）、44.1kHz（16位，2通道）、32kHz（16位，2通道）、32kHz（12位，4通道）。音频采样可以锁定到也可以不锁定到视频帧频上。

516
517

音频数据按帧处理。对于一个锁定的48kHz采样率，每个帧或者包含1600或1602个音频样本（480i系统），或者包含1920个音频样本（576i系统）。对于480i系统，每帧音频样本的数量遵循一个5帧序列：

1600，1602，1602，1602，1602，1600，……

对于一个锁定的32kHz采样率，每个帧或者包含1066或1068个音频样本（480i系统），或者包含1280个音频样本（576i系统）。对于480i系统，每帧音频样本的数量遵循一个15帧序列：

1066，1068，1068，1068，1068，1068，1068，1066，
1068，1068，1068，1068，1068，1068，1068，……

对于未锁定的音频采样，尽管可以指定每帧音频样本的最大值和最小值，但是每帧音频样本的数量不固定。

11.1.2　SMPTE 314M/370M

SMPTE 314M和370M支持更加有限的选项，音频采样要锁定到视频帧频上：

- 25Mbit/s，48kHz（16位，2通道）
- 50Mbit/s，48kHz（16位，4通道）
- 100Mbit/s，48kHz（16位，8通道）

音频数据按帧处理。对于一个锁定的48kHz采样率，每个帧或者包含1600或1602个音频样本（60场/帧制），或者包含1920个音频样本（50场/帧制）。对于60场/帧制，每帧音频样本的数量遵循一个5帧序列：

1600，1602，1602，1602，1602，1600，……

音频容量对于60场/帧制能够满足每帧1620个样本，或者对于50场/帧制能够满足每帧1944个样本。每帧末尾未使用的空间填充的是任意数据。

11.1.3　音频辅助数据（AAUX）

AAUX信息添加到图11-2中所示的散乱的音频数据上。AAUX包组包含一个1字节的包组头和4字节的数据（载荷），结果生成一个5字节的AAUX包组。因为每个视频帧中包含9个AAUX包组，所以AAUX包组的编号为0~8。压缩流中必须包含一个AAUX源（AAUX source，AS）包组和一个AAUX源控制（AAUX source control，ASC）包组。尽管IEC 61834支持许多其他包组格式，但是当前，SMPTE只支持AS和ASC包组。

1. AAUX源（AS）包组

该包组的格式如表11-1所示。

表11-1　AAUX源（AS）包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	0
PC1	LF	1	AF					
PC2	SM	CHN		PA	AM			
PC3	1	ML	50/60	STYPE				
PC4	EF	TC	SMP			QU		
SMPTE 314M/370M	**D7**	**D6**	**D5**	**D4**	**D3**	**D2**	**D1**	**D0**
PC0	0	1	0	1	0	0	0	0
PC1	LF	1	AF					

（续）

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC2	0	CHN		1	AM			
PC3	1	1	50/60	STYPE				
PC4	1	1	SMP			QU		

LF为锁定的音频采样率：0＝锁定到视频，1＝不锁定到视频。

AF为音频帧大小。指定每帧音频样本的数量。

SM为立体声模式：0＝多立体声音频，1＝集总音频（lumped audio）。 518

PA指定记录在CH1（CH3）中的音频信号是否与记录在CH2（CH4）中的音频信号相关：0＝成对通道中的一对，1＝独立通道。

CHN为一个音频块中音频通道的数量：00＝每块1个通道，01＝每块2个通道，10＝保留，11＝保留。

AM指定每个通道上音频信号的内容。

ML为多语言标志：0＝以多语言录制，1＝没有以多语言录制。

50/60表示50场或60场制式：0＝60场制式，1＝50场制式。

STYPE表示对于SMPTE 314M/370M，指定每帧音频块的数量：00000＝2个音频块，00001＝保留，00010＝4个音频块，00011＝8个音频块，00100~11111＝保留。对于IEC 61834，指定视频系统：00000＝标准清晰度，00001＝保留，00010＝高清晰度，00011~11111＝保留。

EF为音频加重标志：0＝打开，1＝关闭。

TC表示加重时间常量：1＝50/15μs，0＝保留。 519

SMP表示音频采样频率：000＝48kHz，001＝44.1kHz，010＝32kHz，011~111＝保留。

QU表示音频量化：000＝16位线性，001＝12位非线性，010＝20位线性，011~111＝保留。

2. AAUX源控制（ASC）包组

该包组的格式如表11-2所示。

表11-2 AAUX源控制（ASC）包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	1
PC1	CGMS		ISR		CMP		SS	
PC2	REC S	REC E	REC M			ICH		
PC3	DRF	SPD						
PC4	1	GEN						

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	1
PC1	CGMS		1	1	1	1	EFC	
PC2	REC S	REC E	FADE S	FADE E	1	1	1	1
PC3	DRF	SPD						
PC4	1	1	1	1	1	1	1	1

CGMS为复制代次管理系统：00＝允许无限制复制，01＝保留，10＝允许复制一次，11＝不允许复制。

ISR为前一输入源：00＝模拟输入，01＝数字输入，10＝保留，11＝没有信息。

CMP为压缩倍数：00＝1倍，01＝2倍，10＝3倍或更大，11＝没有信息。

SS表示源和录制情况：00＝观众受限的加密源并且没有解密录制，01＝观众不受限的加密源并且没有解密录制，10＝观众受限的源或者观众受限的解密源，11＝没有信息。

EFC为音频加重标志：00＝加重关闭，01＝加重打开，10＝保留，11＝保留。

REC S为录制起始点：0＝位于录制起始点，1＝没有位于录制起始点。

REC E为录制结束点：0＝位于录制结束点，1＝没有位于录制结束点。

REC M为录制模式：001＝原始，011＝插入1个CH，100＝插入4个CH，101＝插入2个CH，111＝无效录制。

520

FADE S表示录制起始点的淡入：0＝淡入关闭，1＝淡入打开。

FADE E表示录制结束点的淡出：0＝淡出关闭，1＝淡出打开。

ICH为插入音频通道：001＝CH1，001＝CH2，010＝CH3，011＝CH4，100＝CH1、CH2，101＝CH3、CH4，110＝CH1、CH2、CH3、CH4，111＝没有信息。

DRF为方向标志：0＝反向，1＝正向。

SPD为回放速度。

GEN指示音频源的种类。

11.2　视频

如表11-3所示，IEC 61834对720×480i视频（图3-5）使用4：1：1 YCbCr，对720×576i视频（图3-11）使用4：2：0 YCbCr。

521

表11-3　IEC 61834、SMPTE 314M和SMPTE 370M YCbCr参数

参　数	480i系统	576i系统	720p系统	1080i系统
有效分辨率（Y）	720×480i	720×576i	1280×720p	1920×1080i
帧频	29.97Hz	25Hz	50Hz 59.94Hz	25Hz 29.97Hz
YCbCr采样结构 IEC 61834 SMPTE 314M SMPTE 370M	 4：1：1 4：1：1、4：2：2 	 4：2：0 4：1：1、4：2：2 	 4：2：2	 4：2：2
YCbCr编码格式	均匀量化PCM，每样本8 bit		均匀量化PCM，每样本10 bit	
有效行号	23~262、285~524	23~310、335~622	26~745	21~560、584~1123

SMPTE 314M对25Mbit/s实现的两类视频标准都使用4：1：1 YCbCr（图3-5）。4：2：2 YCbCr（图3-3）适用于50Mbit/s和100Mbit/s的实现。

11.2.1　DCT块

每个帧的Y、Cb和Cr样本分成8×8个块，称为DCT块。除了4：1：1模式期间Cb和Cr最右边的DCT块，每个DCT块传送视频数据的8样本×8行。DCT块的第1、3、5和7排处理第1场，而DCT块的第0、2、4和6排处理第2场。

对于480i系统，每个视频帧有10 800（4：2：2）或者8100（4：1：1）个DCT块。

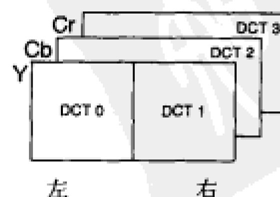


图11-4　4：2：2宏块排列

对于576i系统，每个视频帧有12 960（4：2：2）或者9720（4：1：1、4：2：0）个DCT块。

11.2.2 宏块

如图11-4所示，4：2：2模式中的每个宏块（macroblock）包含4个DCT块。如图11-5和图11-6所示，4：1：1模式和4：2：0模式中的每个宏块包含6个DCT块。

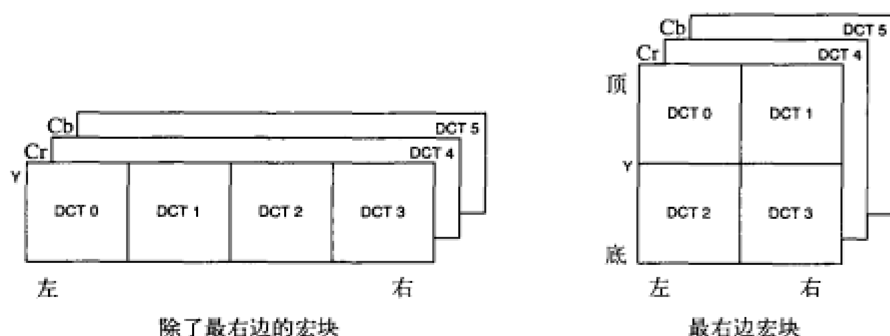


图11-5 4：1：1宏块排列

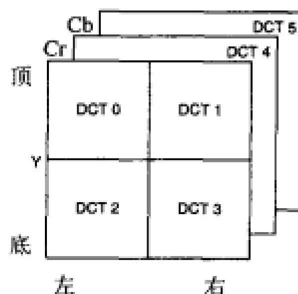


图11-6 4：2：0宏块排列

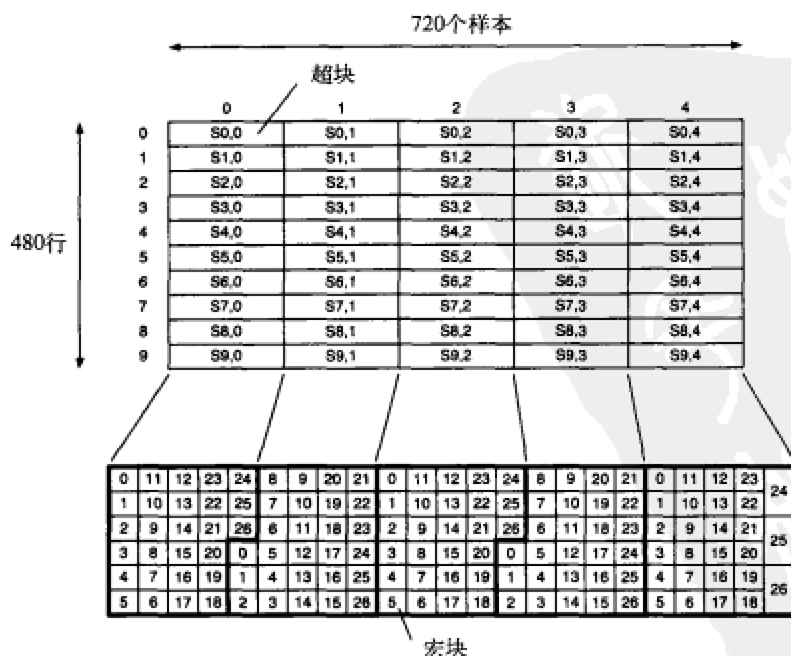


图11-7 超块和宏块之间的关系（4：1：1 YCbCr，720×480i）

对于480i系统，4∶1∶1和4∶2∶2 YCbCr数据的帧的宏块排列分别如图11-7和图11-8所示。每个视频帧有2700（4∶2∶2）或者1350（4∶1∶1）个宏块。

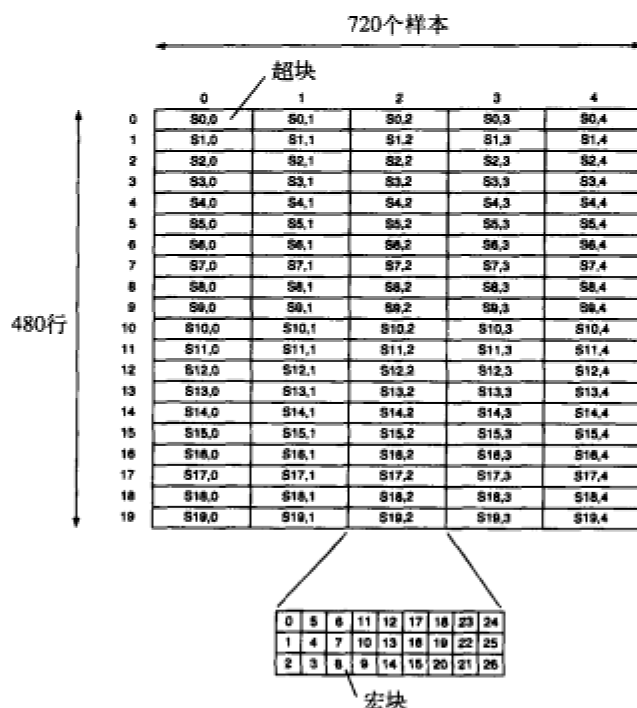


图11-8 超块和宏块之间的关系（4∶2∶2 YCbCr，720×480i）

对于576i系统，4∶2∶0、4∶1∶1和4∶2∶2 YCbCr数据的帧的宏块排列分别如图11-9、图11-10和图11-11所示。每个视频帧有3240（4∶2∶2）或者1620（4∶1∶1、4∶2∶0）个宏块。

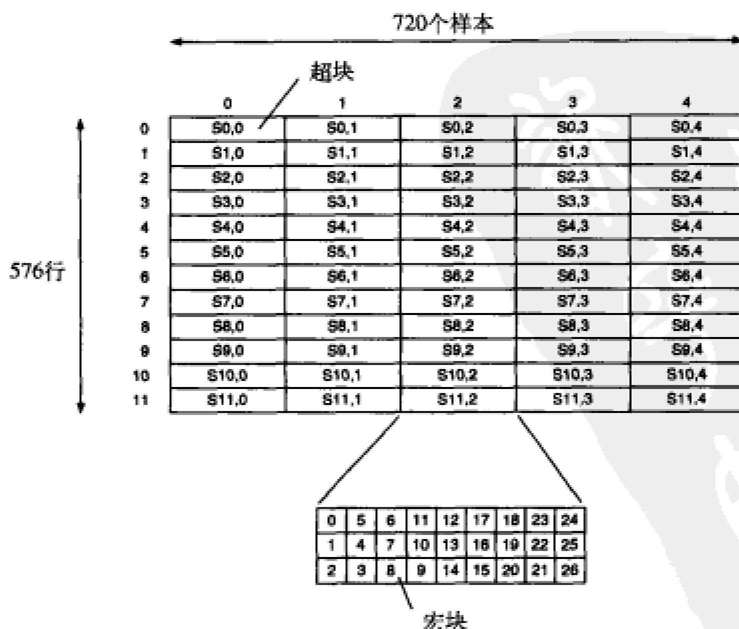


图11-9 超块和宏块之间的关系（4∶2∶0 YCbCr，720×576i）

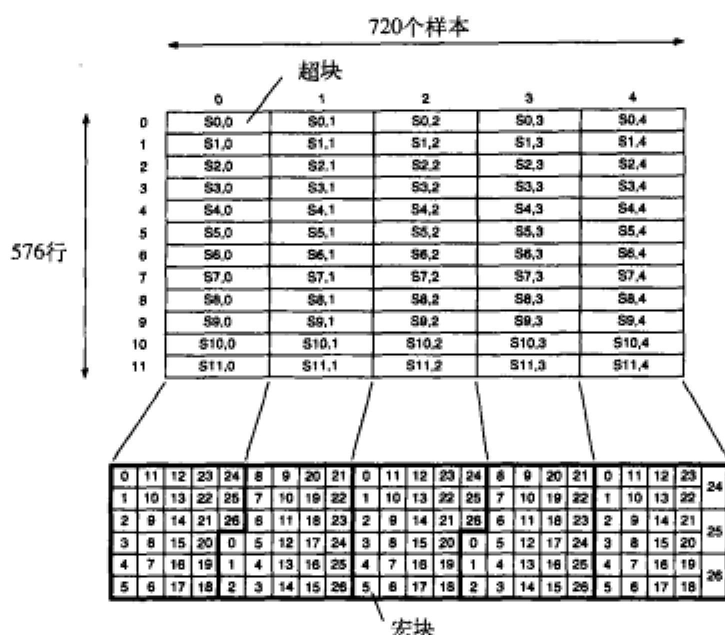


图11-10　超块和宏块之间的关系（4：1：1 YCbCr，720×576i）

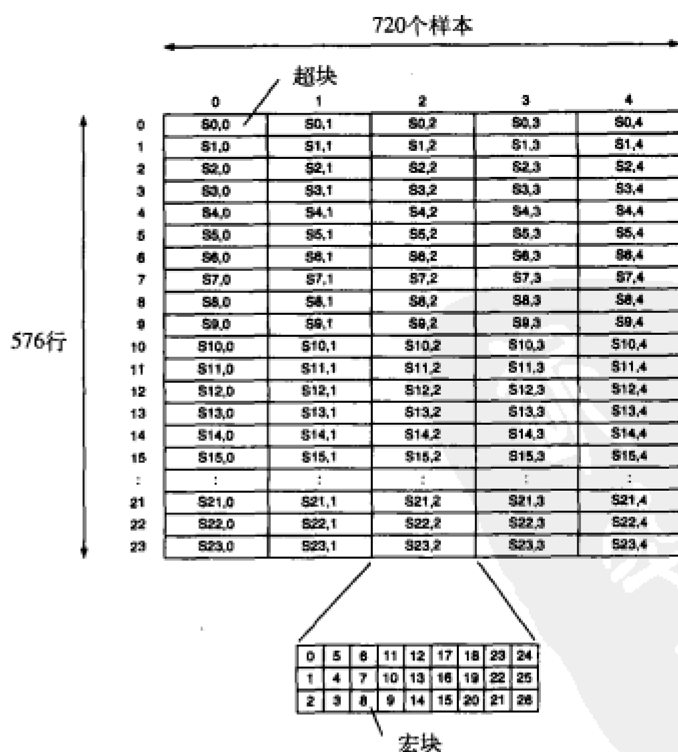


图11-11　超块和宏块之间的关系（4：2：2 YCbCr，720×576i）

11.2.3　超块

每个超块（superblock）包含27个宏块。

对于480i系统，4：1：1和4：2：2 YCbCr数据的帧的超块排列分别如图11-7和图11-8所示。每个视频帧有100（4：2：2）或者50（4：1：1）个超块。

对于576i系统，4：2：0、4：1：1和4：2：2 YCbCr数据的帧的超块排列分别如图11-9、图11-10和图11-11所示。每个视频帧有120（4：2：2）或者60（4：1：1、4：2：0）个宏块。

11.2.4　压缩

像MPEG和H.263一样，DV使用基于DCT的视频压缩。但是，在这种情况下，DCT块由2个场构成，每个场提供4个扫描行中的样本以及8个水平样本。

DCT模式有2种，分别称为8-8-DCT和2-4-8-DCT，根据视频帧两场之间内容的变化程度，压缩转换过程可以选择使用这两种模式。8-8-DCT是普通的8×8 DCT，当两个场之间的相关程度较高（运动很少）时使用。2-4-8-DCT使用2个4×8 DCT（每个场一个），当两个场之间的相关程序较低（运动较多）时使用。关于使用哪个DCT的信息使用一个单独位存储在DC系数中。

DCT系数量化为9位，然后除以一个量化数，从而将一个视频段中的数据量限制到5个压缩宏观中。

522

每个DCT块基于量化噪声（quantization noise）和AC系数的最大绝对值分成4类。2bit的分类号存储在DC系数区域中。

区域号用于量化阶的选择。区域号有4个，它们的选择基于行频和场频。

量化阶由分类号、区域号和量化号（QNO）决定。量化信息传递到视频块的DIF头中。

可变长度的编码将量化的AC系数转换为可变长度的代码。

图11-12和图11-13描述了压缩宏块的排列。

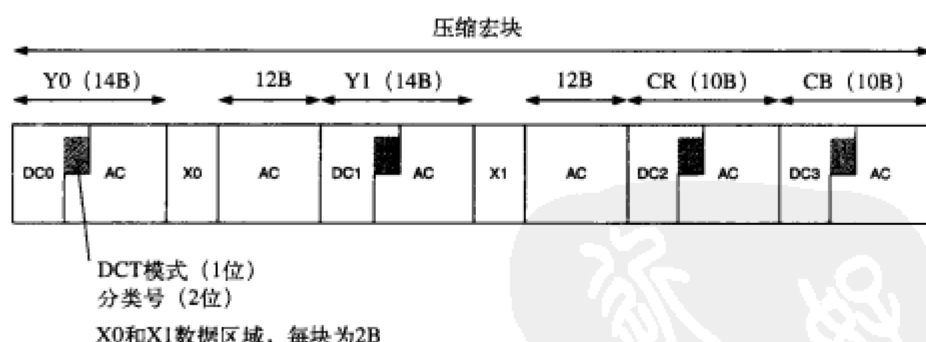


图11-12　4：2：2压缩宏块排列

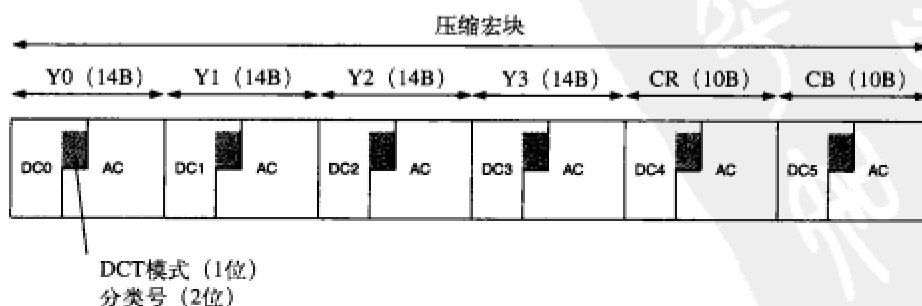


图11-13　4：2：0和4：1：1压缩宏块排列

11.2.5　视频辅助数据（VAUX）

VAUX信息添加到图11-3中所示的散乱的视频数据上。VAUX包组包含一个1字节的包组头和4字节的数据（载荷），结果生成一个5字节的VAUX包组。因为每个视频帧中包含45个VAUX包组，所以VAUX包组的编号从0至44。压缩流中必须包含一个VAUX源（VAUX source，VS）包组和一个VAUX源控制（VAUX source control，VSC）包组。尽管IEC 61834支持许多其他包组格式，但是当前，SMPTE 314M只支持VS和VSC包组。

1. VAUX源（VS）包组

该包组的格式如表11-4所示。

表11-4　VAUX源（VS）包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	TVCH（十位，0~9）				TVCH（个位，0~9）			
PC2	B/W	EN	CLF		TVCH（百位，0~9）			
PC3	SRC		50/60	STYPE				
PC4	TUN							

SMPTE 314M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	1	1	1	1	1	1	1	1
PC2	B/W	EN	CLF		1	1	1	1
PC3	1	1	50/60	STYPE				
PC4	VISC							

SMPTE 370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	1	1	1	1	1	1	1	1
PC2	1	1	1	1	1	1	1	1
PC3	1	1	50/60	STYPE				
PC4	0	1	1	1	1	1	1	1

TVCH为电视频道号，0~999。值0xEEE为预录制磁带或者线输入保留。值0xFFF为“没有信息”保留。

B/W为黑白标志：0＝黑白视频，1＝彩色视频。

523

EN为CLF有效标志：0＝CLF有效，1＝CLF无效。

CLF为彩色帧标识码。对于480i系统：00＝彩色帧A，01＝彩色帧B，10＝保留，11＝保留。对于576i系统：00＝第1、2场，01＝第3、4场，10＝第5、6场，11＝第7、8场。

SRC定义视频信号的输入源。

50/60同AAUX。

STYPE表示对于SMPTE 314M，指定视频信号类型：00000＝4：1：1压缩，00001~00011＝保留，00100＝4：2：2压缩，00101~11111＝保留。对于SMPTE 370M，指定视频信号类型：00000~10011＝保留，10100＝1080i30或1080i25，10101＝1035i30，10110＝保留，10111＝保留，11000＝720p60或720p50，11001~11111＝保留。对于IEC 61834，指定视频制式：00000＝标准清晰度，00001＝保留，00010＝高清晰度，00011~11111＝保留。

TUN为调谐器类别，包含3位区域号和5位卫星号。11111111表示没有信息可用。

VISC：10001000＝－180……00000000＝0……01111000＝180，01111111＝没有信息，其他值＝保留。

2. VAUX源控制（VSC）包组

该包组的格式如表11-5所示。

表11-5 VAUX源控制（VSC）包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	1
PC1	CGMS		ISR		CMP		SS	
PC2	REC S	1	REC M		1	DISP		
PC3	FF	FS	FC	IL	SF	SC	BCS	
PC4	1	GEN						

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	1
PC1	CGMS		1	1	1	1	1	1
PC2	1	1	0	0	1	DISP		
PC3	FF	FS	FC	IL	1	1	0	0
PC4	1	1	1	1	1	1	1	1

CGMS同AAUX。

ISR同AAUX。

CMP同AAUX。

SS同AAUX。

524～532 REC S同AAUX。

REC M表示录制模式：00＝原始，01＝保留，10＝插入，11＝无效录制。

BCS为广播制式，指示带有DISP的显示格式的信息类型：00＝类型0（IEC 61880、CEA-608），01＝类型1（ETS 300 294），10＝保留，11＝保留。

DISP为宽高比信息。

FF为帧/场标志。指示两个场/帧是按顺序输出，还是在一/两个帧周期内输出一个场两次：0＝一个场/帧输出两次，1＝两个场/帧按顺序输出。

533 FS为第一/第二标志。指示在场/帧1周期内输出哪个场/帧：0＝场/帧2，1＝场/帧1。

FC为帧变化标志。指示当前帧的图像是否与刚才前一帧的图像相同：0＝相同图像，1＝不同图像。

IL为隔行标志，指示构成一帧的两场的数据是隔行的还是逐行的：0＝逐行，1＝隔行或者无法识别。

SF为静止场图像标志，指示一帧内两场之间的时间差：0＝0秒，1＝1001/60或者50秒。

SC为照相机图像标志：0＝照相机图像，1＝非照相机图像。

GEN指示视频源的类别。

11.3 数字接口

IEC 61834和SMPTE 314M都指定了通用数字接口的数据格式。例如，该数据格式可以通过IEEE 1394或者SDTI发送。图11-14描述了帧数据结构。

每个720×480i 4∶1∶1 YCbCr帧压缩为103 950B。包含开销和音频时数据量增加到120 000B。

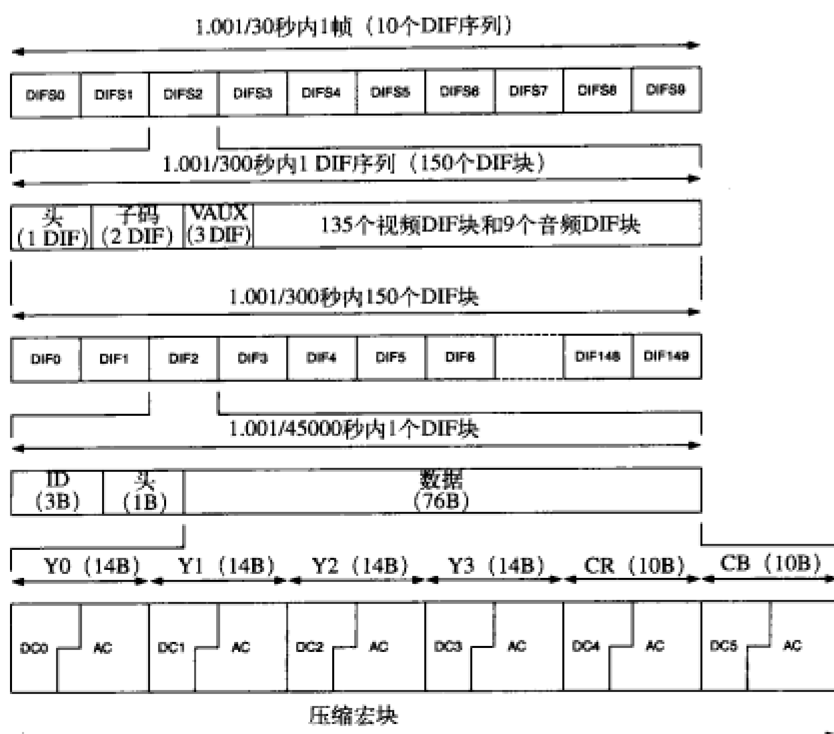


图11-14 25Mbit/s 4：1：1 YCbCr 720×480i系统的包格式

压缩720×480i帧分成10个DIF（data in frame，帧中数据）序列。每个DIF序列包含150个DIF块，每个DIF块80字节，DIF序列的分配情况如下所示：

- ❑ 135个DIF块用于视频；
- ❑ 9个DIF块用于音频；
- ❑ 6个DIF块用于头、子码和视频辅助（VAUX）信息。

图11-14详细描述了DIF序列的结构。每个视频DIF块包含80字节压缩宏块数据：

- ❑ 3字节用于DIF块ID信息；
- ❑ 1字节用于包含量化号（QNO）和块状态（STA）的头；
- ❑ Y0、Y1、Y2和Y3各使用14B；
- ❑ Cb和Cr各使用10B。

720×576i帧可以使用4：2：0 YCbCr格式（IEC 61834）或者4：1：1 YCbCr格式（SMPTE 314M），而且需要12个DIF序列。每个720×576i帧压缩为124 740字节。包含开销和音频时数据量增加到144 000B，要求300个包进行传送。

注意通过接口传送的数据组织与实际DV录制格式有所不同，这是因为数字传输不需要错误校正。另外，尽管视频块按图11-15中的顺序编号，但是该序列并不对应于视频数据块从左至右、从上到下的传输顺序。压缩宏块散乱传送是为了使错误的影响最小，有助于错误隐蔽。音频数据也是散乱传送的。数据按与录制相同的散乱顺序传送。

为了描述视频的散乱性，DV视频帧被组织成超块，每个超块由27个压缩宏块构成，如图11-7~图11-11所示。5个超块为一组（从每个超块列中取一个超块），构成一个DIF序列。表11-6和表11-7描述了DIF块的传输顺序。 534

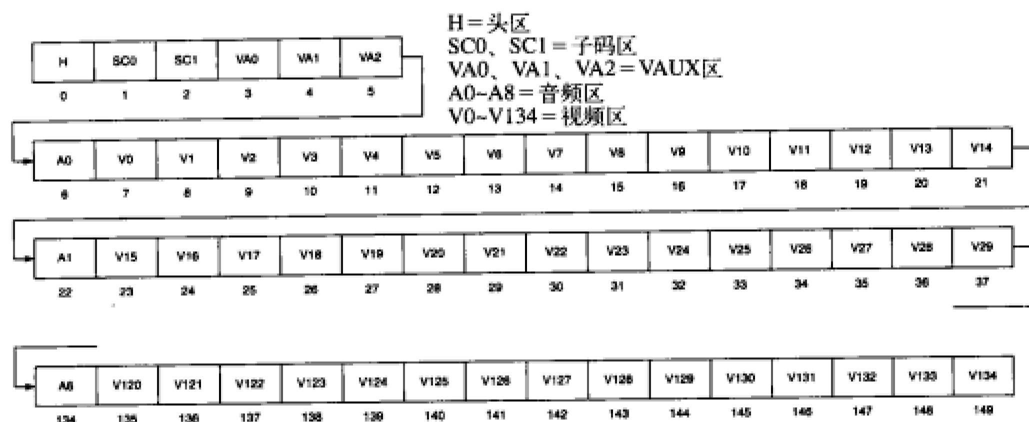


图11-15　DIF序列细节

表11-6　25Mbit/s（4：1：1或4：2：0 YCbCr）的视频DIF块和压缩宏块

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
0	0	2，2	0	⋮			
	1	6，1	0	$n-1$	0	1，2	0
	2	8，3	0		1	5，1	0
	3	0，0	0		2	7，3	0
	4	4，4	0		3	$n-1$，0	0
	⋮				4	3，4	0
	133	0，0	26		⋮		
	134	4，4	26		133	$n-1$，0	26
1	0	3，2	0		134	3，4	26
	1	7，1	0				
	2	9，3	0				
	3	1，0	0				
	4	5，4	0				
	⋮						
	133	1，0	26				
	134	5，4	26				

注：对于480i系统$n=10$，对于576i系统$n=12$。

表11-7　50Mbit/s（4：2：2 YCbCr）的视频DIF块和压缩宏块

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
0	0，0	4，2	0	⋮			
	0，1	5，2	0	$n-1$	0，0	2，2	0
	1，0	12，1	0		0，1	3，2	0
	1，1	13，1	0		1，0	10，1	0
	2，0	16，3	0		1，1	11，1	0
	⋮				2，0	14，3	0
	134，0	8，4	26		⋮		
	134，1	9，4	26		134，0	6，4	26

（续）

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
1	0，0	6，2	0	$n-1$	134，1	7，4	26
	0，1	7，2	0				
	1，0	14，1	0				
	1，1	15，1	0				
	2，0	18，3	0				
	⋮						
	134，0	10，4	26				
	134，1	11，4	26				

注：对于480i系统$n=10$，对于576i系统$n=12$。

对于50Mbit/s的SMPTE 314M格式，每个压缩720×480i帧或720×576i帧分成2个通道。每个通道使用10个（480i系统）或者12个（576i系统）DIF序列。

11.3.1 IEEE 1394

使用IEEE 1394接口传送DV信息的内容在第6章中讨论。

11.3.2 SDTI

SDTI的一般概念在第6章中讨论。

1. SMPTE 314M数据

SMPTE 221M详细描述了如何通过SDTI传送SMPTE 314M DV数据。

2. IEC 61834数据

SMPTE 222M详细描述了如何通过SDTI传送IEC 61834 DV数据。

535

11.4 100Mbit/s DV不同之处

100Mbit/s SMPTE 370M格式支持1920×1080i和1280×720p源。

1920×1080i源缩放为1280×1080i。1280×720p源缩放为960×720p。所用采样为4：2：2 YCbCr采样。

每个压缩帧分成4个通道。每个通道使用10个（1080i30或者720p60系统）或者12个（1080i25或者720p50系统）DIF序列。

11.5 HDV格式

HDV格式由佳能、夏普、索尼和JVC共同开发，HDV支持录制1920×1080和1280×720内容，使用的是标准的DV磁带。HDV格式基于25Mbit/s MPEG-2分组基本流和19Mbit/s MPEG-2传输流，视频压缩使用MPEG-2，音频压缩使用MPEG-1 Layer Ⅱ。

11.6 AVCHD格式

AVCHD格式由松下和索尼公司共同开发，AVCHD支持录制1920×1080、1280×720、720×576和720×480内容，使用8cm DVD-RW、8cm BD-R/RE、SD存储卡或者HDD来代替磁带。AVCHD格式基于24Mbit/s MPEG-2传输流，视频压缩使用MPEG-4.10（H.264），音频压缩使用

[536] Dolby® Digital或者LPCM。

参考文献

1. IEC 61834-1, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 1: General specifications.*
2. IEC 61834-2, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 2: SD format for 525-60 and 625-50 systems.*
3. [537] IEC 61834-4, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 4: Pack header table and contents.*
4. SMPTE 314M-2005, *Television—Data Structure for DV-Based Audio, Data and Compressed Video-25 and 50 Mbps.*
5. SMPTE 321M-2002, *Television—Data Stream Format for the Exchange of DVBased Audio, Data and Compressed Video Over a Serial Data Transport Interface.*
6. SMPTE 322M-2004, *Television—Format for Transmission of DV Compressed Video, Audio and Data Over a Serial Data Transport Interface.*
7. SMPTE 370M-2006, *Television—Data Structure for DV-Based Audio, Data and Compressed Video at 100 Mb/s 1080/60i, 1080/50i, 720-60p, 720/50p.* [538]



第12章 MPEG-1

MPEG-1视音频压缩标准是为存储和分发视频和音频数据而开发的，其特征包括随机存取、快进和快退播放等功能，是早期视频CD（即VCD）的基础标准。

在CD时代，人们根据可获得的媒体来定义信道带宽和图像分辨率，其目标是采用比特率为1.416Mbit/s（其中1.15Mbit/s用于视频）的标准光盘来播放数字视频和音频。

作为一个ISO标准，MPEG-1（ISO/IEC 11172）由以下6部分组成：

系统	ISO/IEC 11172-1
视频	ISO/IEC 11172-2
音频	ISO/IEC 11172-3
低比特率音频	ISO/IEC 13818-3
一致性测试	ISO/IEC 11172-4
仿真软件	ISO/IEC 11172-5

MPEG-1的比特流隐式地定义了解压缩算法。而压缩算法则由各厂商独立开发，并允许在国际标准范围内获得所属权益。

12.1 MPEG和JPEG

JPEG（ISO/IEC 10918）是为压缩静态连续色调（continuous-tone）的灰度图像或彩色图像而设计的。JPEG不善于表现只有黑白两种颜色的二值图，此外，如果要处理伪彩色图像，必须先将图像转换成未上色（unmapped color）的形式。JPEG图像可以针对任意分辨率和色彩空间，处理方法包括无损或者有损两种类型。

作为一种通用的标准，JPEG特色丰富，功能强大。通过调整各种各样的参数，压缩后图像大小与重建图像质量之间可以在一个较大的范围内获得折中。图像质量可以从“浏览级”（100：1的压缩比）到“无失真级”（约3：1的压缩比）之间变化。一般来说，源图像按10：1到20：1的比率压缩，才能看出重建图像与源图像之间存在视觉差异。

JPEG并不是采用单一算法，而是一系列的压缩算法，包含4种方式，分别应用于不同领域。其中最为人们熟知的有损压缩算法是顺序DCT，顺序DCT可能采用Huffman编码（基本JPEG）或算术编码。图像解码的时候，顺序从左到右，自上而下。

另一种有损压缩算法是累进DCT，需要对一幅图像进行多次扫描。图像解码的时候，首先得到整幅图像的粗糙效果，随后逐渐清晰，直至解码完成。这样做有利于图像数据库浏览等应用场合。累进DCT可能采用光谱选择或者连续逼近两个选项中的一个，或者两个选项都使用。为快速得到一幅图像，光谱选择选项首先对低频DCT系数编码，再逐渐增加细节，接着对高频系数编码。连续逼近选项首先对重要的DCT系数编码，再对次要的DCT系数编码。

层次化模式将图像表示为多种分辨率，如可能有512×512、1024×1024和2048×2048等几种版本的图像。对较高分辨率图像与其相邻最近的低分辨率图像之间的差异进行编码，用编码结果来表征较高分辨率图像，从而使得编码所需的比特数比两者独立存储时要少很多。当然，这样做也使得总的比特数比仅仅存储最高分辨率图像多了不少。此外，层次序列中的单个图像可能会根

据需要按照累进模式编码。

与变换域不同的是，在像素域还有一种无失真的空间压缩算法。该算法利用一个像素点邻近的三个像素点值估算该点的值，然后再减去该点的实际值，将差值采用Huffman或者算术编码方法进行无失真编码。无失真编码操作能达到约2：1的压缩比。

考虑到视频实际上是一系列静态图像，且基本JPEG编解码容易实现，因此有人提出利用基本JPEG压缩实时视频（这种方法也被称为运动JPEG或MJPEG）。但是，这种方法没有利用帧与帧之间的冗余来提高压缩能力，而MPEG利用了。

除此之外，也许最重要的一个问题是，JPEG是对称的，即编码器和解码器的开销大致相似。与此不同的是，*MPEG是以对视频进行一次压缩、而在各种平台多次播放为主要设计目的*，因此，为最小化MPEG硬件解码器开销，MPEG被设计成为不对称的，即编码过程所需计算量是解码过程的100倍。

因为*MPEG面向特定应用，所以其硬件仅支持一些特定的分辨率*。同样，MPEG仅支持一种8位采样的颜色空间YCbCr。此外，MPEG仅在有限的压缩率范围中优化应用。

如果是将捕获视频用于编辑，可以采用基本JPEG或者单纯I帧（内码帧）MPEG将内容实时压缩到盘片。采用JPEG方法，需要系统为SIF（标准输入格式）分辨率提供4Mbit/s的传输和硬盘访问能力。一旦编辑完毕，结果可转换成MPEG以达到最大程度的压缩。

12.2　质量问题

在比特率约3~4Mbit/s的时候，视频传输采用MPEG-1可达到“广播级质量”。然而，对于复杂的时空场景变化（如体育运动节目）序列，由于MPEG-1基于帧的处理，导致传输比特率需求提高至5~6Mbit/s。MPEG-2支持基于场的处理，因而允许“广播级质量”的比特传输率达到4~6Mbit/s。

540

影响MPEG压缩视频质量的因素包括原始视频源的分辨率、压缩后信道带宽容许的比特率和运动估计效果。

影响压缩视频质量的一个重要因素是原始视频源的分辨率，如果原始分辨率过低，细节将普遍缺失。

运动估计效果决定运动失真情况，也就是当运动开始或者运动量超过一定限度时产生的视频质量下降状况。差的运动估计会导致视频质量效果整体下降。

最重要的是，比特率（信道带宽）越高，可以传输的信息量就会越大，此时运动失真减少，或者能够显示更高分辨率的图像。一般来说，降低比特率不会导致解码视频质量下降得太厉害。而一旦比特率下降到一定限度，视频质量将迅速下降，8×8的方块将变得清晰可见。

12.3　音频概述

MPEG-1采用一系列共三个层次的音频编码方案，称为层I、层II和层III，计算复杂度和声音质量逐层提高。这三个层是分等级的：层III解码器处理层I、层II和层III；层II解码器只处理层I、层II；层I仅处理层I。所有层都支持采用16kHz、22.05kHz、24kHz、32kHz、44.1kHz或48kHz采样频率的16位音频。

对于每一层，其比特流格式和解码器是指定的。而编码器没有指定，允许后续改进。所有层工作在相似的比特率下。

- 层I：32~448kbit/s。
- 层II：8~384kbit/s。

❑ 层Ⅲ：8~320kbit/s。

两个声道支持四种操作模式：正常立体声，联合（增强以及中间）立体声，双通道单声道，单通道单声道。在正常立体声模式下，一个信道传送左声道音频信号，另一个信道传送右声道音频信号。在增强立体声模式（所有层都支持）下，2kHz以上的高频子带将合成一个信号。在这种模式下，立体声像得以保持，但是仅传送时域包络。在中间立体声编码（仅层3支持）模式下，一个信道负责传送和信号（左声道＋右声道），另一个传送差信号（左声道－右声道）。此外，还支持传送预加重（pre-emphasis）、版权标记以及原创/复制说明等信息。

12.3.1 声音质量

为确定对于特定应用，具体应采用哪一层方案，需要考察各层可达到的比特率（因为各层都支持在一定比特率下，声音质量下降最小）。

层I，作为层Ⅱ的简化版本，目标比特率是每信道192kbit/s或更高。

层Ⅱ等同于MUSICAM，目标比特率为每信道128kbit/s，用于在声音质量与编码复杂度之间的折中。层Ⅱ特别适合于比特率为每信道96~128kbit/s。

层Ⅲ（即大家熟知的mp3）融合MUSICAM和ASPEC的思想，速率能达到每信道64kbit/s。层Ⅲ格式定义了一系列高级特性，只为完成一个共同的目标：在相对低的码率下尽可能保证声音质量。

541

12.3.2 背景理论

所有层均采用基于心理声学原则（特别是掩蔽效应）的编码策略。例如，根据掩蔽效应，收听到的响度较高的频率成分会掩盖响度较低的频率成分。

假设有一个强音调频率为1000Hz，另有一个1100Hz的音调，强度比前者低18dB。则1100Hz的音调将不会被听到，它被1000Hz的音调掩蔽了。然而，一个频率为2000Hz、强度比1000Hz的音调低18dB的音调则可以被听到。2000Hz的音调要想被1000Hz的音调掩蔽，则需要降低约45dB。任何相对低的频率都会被高频掩蔽，另外，距离声源越远，掩蔽效应越不明显。

因此，可以做出相对能量与掩蔽频率（当前掩蔽）之间的曲线图。如果响度变化显著（30~40dB），掩蔽效应可以在一个强音之前发生（预掩蔽）或之后发生（后掩蔽）。其原因是大脑需要反应时间。预掩蔽时间大概是2~5ms，而后掩蔽可以持续到100ms。

调整噪声基底可以减少需要的数据量，使得压缩效果更好。CD采用16位分辨率，可达到近96dB的信噪比（SNR），这正好符合人耳听力域的动态范围（也就是大多数人在播放静音时不会听到噪声）。如果采用8位分辨率，在音乐或者语声的间隙会出现明显的噪声。然而，由于隐蔽效应，高声带中的噪声不会被引起注意，也就是在声音较响的时候，可以提高噪声基底，因为噪声会被掩蔽。

对于立体声信号，信道间经常有冗余，所有层都可以采用联合立体声模式减少这些立体声效果，联合立体声是层Ⅲ中采用的最灵活的方法。

12.4 视频编码层

MPEG-1允许在每秒60帧的逐行扫描方式下，分辨率最高可达到4095×4095。正如很多人认为的，MPEG-1是约束参数比特流（Constrained Parameters Bitstream，CPB）的子集。CPB是包括采样和比特率参数的有限集合，这些参数用来标准化缓存大小和存储带宽，使得编解码器有名义上确保的互操作，同时解决最大可能的应用范围。不能够处理这些参数的设备不能认为是真正的

MPEG-1。表12-1列出了一些约束参数。

CPB限制视频为396个宏块（共101 376像素）。因此，MPEG-1视频通常采用分辨率为352×240p或者352×288p的SIF格式编码。在编码中，分辨率为704×480i或704×576i的原始BT.601按比例缩减成SIF格式分辨率。这通常是通过忽略场2，在水平方向按比例缩减场1来实现的。在解码中，SIF分辨率被按比例提高成704×480i或者704×576i分辨率。注意，为保证Y分量数值是16的倍数，某些完整的有效扫描线和扫描线上的像素会被抛弃。表12-2列出了比较常见的MPEG-1分辨率。

542

表12-1 MPEG-1中的一些约束参数

线分辨率	≤768个样本点
场分辨率	≤576条扫描线
图像面积	≤396个宏块
像素速率	≤每秒396×25个宏块
图像速率	≤每秒30帧
比特率	≤1.856Mbit/s

表12-2 常见MPEG-1分辨率

分 辨 率	帧/秒
352×240p	29.97
352×240p	23.976
352×288p	25
320×240p[1]	29.97
384×288p[1]	25

注：1. 方块像素格式

编码视频速率被控制在1.856Mbit/s上，但是，在某些带宽为6Mbit/s或更高的应用中，比特率是个常常被舍弃的参数。

MPEG-1视频数据采用4：2：0的YCbCr格式，如图3-7所示。

12.4.1 隔行视频

MPEG-1曾被用作处理逐行（即非隔行）视频。在早期，为尽力提高视频质量，人们设计了几种方法来充分利用隔行图像的两场数据。

例如，将两场合为一帧704×480p或704×576p分辨率的图像，并对此图像进行编码。在解码的时候，再将两场分开。然而，这样将导致运动对象产生运动效应，在两场中有些错位。而对两场独立编码虽然避免了运动效应，但是没利用两场中的冗余信息，降低了压缩率。

除此之外，还有其他一些处理隔行视频的方法，因此，MPEG-2定义了一个处理隔行视频的标准（这一部分将在第13章中详细介绍）。

12.4.2 编码预处理

在进行MPEG编码之前，通过对视频流预处理，可以得到更好的图像。

在对特定图像编码的时候，为避免严重失真，可以对整幅图像（或者指定的区域）采取预滤波。在压缩处理前的预滤波与模/数转换前的抗混叠（anti-alias）滤波类似。预滤波考虑纹理样式、运动和边缘特点，可应用到图像、条带、宏块或者块等不同层次上。

对于快速或随机运动较少以及光照条件较好的场景，MPEG编码效果最好。此时，前景光应该清晰，背景光模糊。前景对比度和细节应当正常，而背景对比度最好能比较低，且边缘较为模糊。编辑工具通常会允许你预处理潜在问题区域。

MPEG-1规范中有将BT.601按比例缩减到SIF分辨率的滤波器样例。在实例中，场2被忽略，一半场分辨率被抛弃，并且采用抽取滤波器（抽取因子为2）减少余下扫描线的线分辨率。同样，对于Cb、Cr分量也必须进行适当的抽取。

按比例缩减成SIF分辨率前，对原始图像进行去隔行（deinterlacing）操作，可以得到更好的

视频质量。当处理宏块（这个名词稍后定义）的时候，如果两场间的宏块差别较小，可将两宏块求平均，产生一个新的宏块。另外，采用来自具有同奇偶性（same parity）场的宏块区域，可以避免运动效应的产生。

12.4.3 编码帧类型

编码帧分为4类。I（intra）帧（约1比特/像素）是作为独立静态图像编码的帧。I帧提供在视频流中随机存取的随机存取点。通常，在1秒钟视频中约出现2次I帧。此外，在场景发生突变的地方也应该采用I帧。 543

P（predicted）帧（约0.1比特/像素）是由与其相邻的前一个I帧或P帧预测而得到的帧（即所谓的前向预测处理），如图12-1所示。通过运动补偿技术，P帧比I帧压缩得更小，而且能够作为B帧和接下来的P帧的参考帧。

B（bi-directional）帧（约0.015比特/像素）是采用最邻近的前后的两个I帧或P帧作为参考帧、进行双向预测得到的帧，如图12-1所示。B帧有最高的压缩效率，并利用前后两帧的均值减少了噪声。通常，I帧或P帧之间有2个B帧。

D（DC）帧是仅包括DCT变换中的直流系数成分、按照独立静态图像编码的帧。D帧不能混在其他帧当中，而且很少使用。

图像组（Group Of Picture，GOP）是用于辅助随机存取和编辑的一帧或多帧编码帧序列。在编码过程中GOP值是可以设置的，GOP值越小，对运动的响应效果越好（因为I帧靠得越近），但压缩效率却下降了。 544

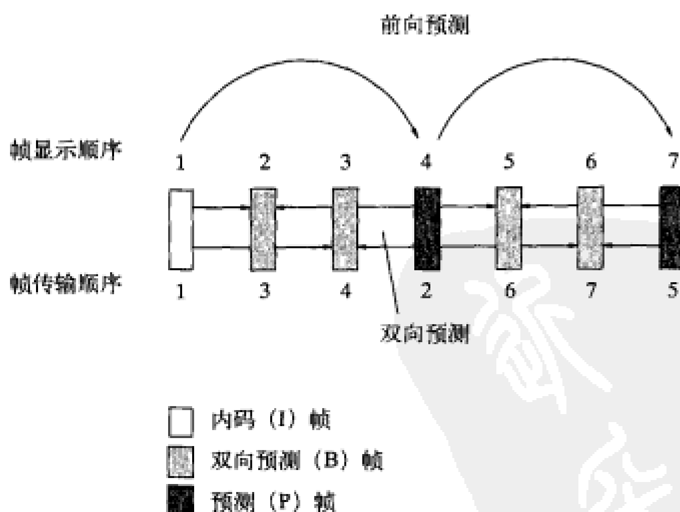


图12-1 MPEG-1中的I帧、P帧和B帧。不按照显示顺序传输的帧序列会使插值过程变得复杂，因此需要MPEG解码器对其重新排序。箭头表示帧之间的依赖关系

在编码比特流中，一个图像组必须以I帧作为开始，随后可以任意顺序接任意个I帧、P帧或B帧。在显示顺序中，图像组必须以I帧或B帧开始，以I帧或P帧结束。因此，最小的图像组是单个I帧，而最大图像组的大小是无限的。

最初，每个图像组不依赖其他图像组进行编码和显示。然而，这种情况只有没有B帧在I帧前才有可能，否则只能进行逆向运动补偿。于是形成开放和封闭两种图像组格式。封闭图像组不需要之前的图像组做运动补偿就能完成解码，而开放图像组是需要的。

12.4.4 运动补偿

运动补偿用于消除帧间的时间冗余，提高对P帧和B帧的压缩效果。运动补偿主要在宏块层（稍后介绍）发挥作用。

运动补偿技术原理基于以下事实：在几幅大致相同图像的序列中，多数对象基本处在相同的位置，其余对象移动了一小段距离。这些运动被描述成一个二维运动矢量，这个二维运动矢量用来指定从前一个解码帧中的哪个位置获得一个宏块，该宏块用来预测当前宏块的像素值。

545

在采用运动补偿压缩宏块之后，待编码宏块和参考宏块之间存在空间差异（运动矢量）和内容差异（误差项）。

注意，有可能出现一个场景的信息不能从前一个场景预测出来的情况，例如打开一个门，先前的场景是没有包含门后区域的细节信息的。在这些情况下，当P帧中的宏块不能表征运动补偿的时候，将采用与I帧中的宏块编码方式相同的方式进行编码（采用帧间编码模式）。

B帧中的宏块采用最邻近的前后I帧或P帧作为参照，于是有四种编码方式：

- ❑ 帧内编码，没有运动补偿；
- ❑ 前向预测，最近的前一个I帧或P帧作为参考；
- ❑ 后向预测，最近的后一个I帧或P帧作为参考；
- ❑ 双向预测，两帧都用作参考——最近的前一个I帧或P帧，以及后一个I帧或P帧。

后向预测用于预测前一帧没有覆盖的区域。

12.4.5 I帧

图像块和预测误差块有很高的空间冗余。为提高压缩率，去除冗余，采用了以下几个步骤，在解码的时候，采用相反的步骤可以恢复数据。

1. 宏块

宏块（如图7-55所示）由16像素×16行的Y分量和对应的8像素×8行的Cb、Cr分量组成。

块是8像素×8行的Y、Cb、Cr值的集合。注意Y块大小为图像（即相应的Cb、Cr块）大小的四分之一。因此，一个宏块包含4个Y块、1个Cb块和一个Cr块，如图12-2所示。

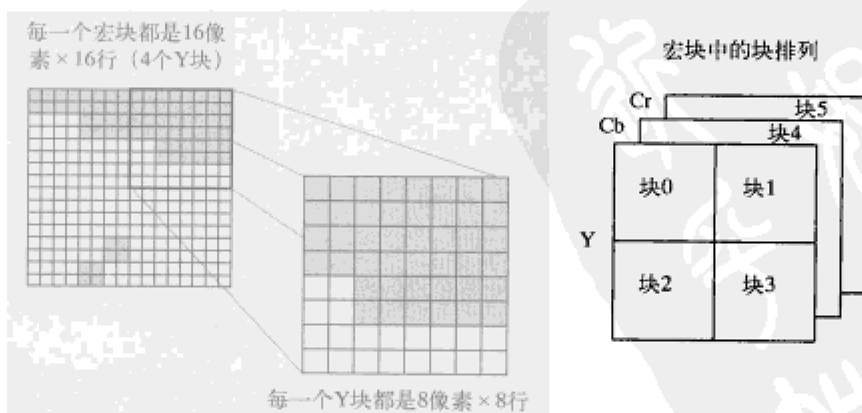


图12-2 MPEG-1宏块和块

I帧内有两种宏块类型，两者都采用帧内编码的，如表12-9所示。一种（称为intra-d）采用当前的量化器刻度，另一种（称为intra-q）采用新的量化器刻度。

如果宏块类型为intra-q，宏块首部指定一个量化器刻度因子，长度为5位。解码器用它从传送

的量化系数中计算DCT系数。量化器刻度因子可以从1到31，但不允许为0。

如果宏块类型为intra-d，将不传送量化器刻度，解码器使用当前的刻度。

2. DCT

每个输入像素或预测误差项的8×8块采用8×8的DCT（离散余弦变换）处理，得到8×8的水平和垂直方向的频率系数块，如图7-56所示。

输入像素值范围是0~255，得到范围为0~2040的DC系数和−1000~+1000的AC系数。

546

3. 量化

频率系数的8×8块是均匀量化的，取值有限。量化器步长刻度由量化矩阵和量化器刻度决定，不同的系数对应不同的量化器步长刻度，宏块与宏块间的量化器步长也有可能发生变化。

DC系数的量化器步长大小固定为8。DC量化系数由DC系数除以8取整得到。AC系数由帧内量化矩阵量化得到。

4. Z形扫描

如图7-57所示，Z形扫描从DC分量开始，按照频率递增的顺序将量化频率系数排列线性流，这将产生长游程的零系数。

5. 量化DC系数编码

DC系数量化后，将采用无失真编码方式对其进行编码。

宏块中的Y块编码按照图12-2所示的顺序进行。块4的DC值为下一个宏块中块1的DC值做预测。在每个条带的开头，DC预测值被置为1024。

每个Cb、Cr块的值采用先前宏块的对应块的DC值进行预测编码，在每个条带开头，DC预测值被置为1024。

DCT变换后的DC差分值按照绝对值形式组织，如表12-16所示。[size]项指定定义Level所需的附加比特数目，[size]用一个可变长编码传送，由于统计结果的不同，这个可变长编码值在Y和CbCr块中也是不同的。例如，size为4表示需要为4个附加比特。

547

解码器按照相反的过程反量化DC系数。

6. 量化AC系数编码

AC系数量化后，将按照图7-57所示的Z形顺序扫描，采用游程（run-length）和非零值（level）编码。如图7-57所示，扫描从位置1开始，因为在位置0的DC系数单独编码。

游程和非零值采用如表12-18所示方式编码。s位表示该非零值的符号，0为正，1为负。

对于表12-18中没有给出的游程－非零值组合的情况，引入了一个转义序列，由一个转义码（ESC）后面接着如表12-19所示的游程和非零值编码。

对最后一个DCT系数编码完成后，添加一个EOB符号，指示解码器在这个8×8块中已经没有其他量化系数了。

12.4.6 P帧

1. 宏块

根据运动补偿产生的复杂性程度不同，P帧分为8种类型的宏块，如表12-10所示。

不编码宏块（skipped macroblock）为零运动矢量的预测宏块。因此，不需要修正；解码器复制先前帧的不编码宏块放在当前帧。不编码宏块的优点是它们几乎不需要传输，它们不需要编码，而是通过指明宏块地址增量把它们跳过。

如果表12-10中的[macroblock quant]项为1，那么需要传送量化器刻度。对剩下的宏块类型，采用先前的量化器刻度进行DCT修正编码。

如果表12-10中的[motion forward]项为1，水平和垂直前向运动矢量应当相继传输。

如果表12-10中的[coded pattern]项为1，6位编码块模式（coded block pattern）将作为变长编码传输。它会指示解码器宏块的6个块中哪些已经编码（设为1），哪些没有编码（设为0）。表12-14列出63种可能组合对应的码字。如果没有块被编码，就没有码字。编码块模式由宏块类型决定。对于I帧中的宏块以及P帧和B帧中的帧内编码宏块，编码块模式将不需要传输，且假定为63（表示所有块都已被编码）。

为确定采用何种类型的宏块，编码器通常需要做出一系列决策，如图12-3所示。

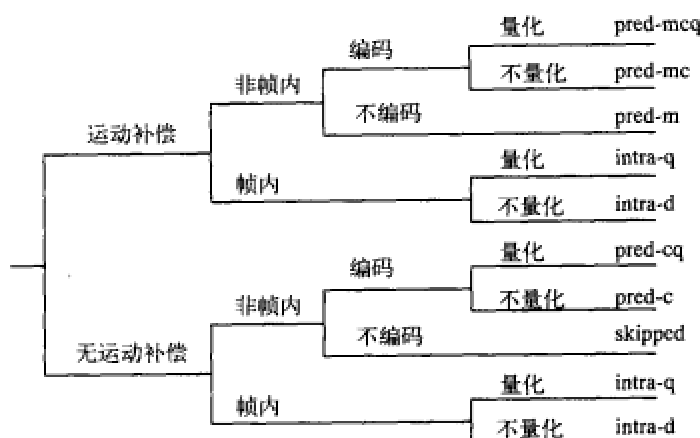


图12-3　MPEG-1 P帧宏块类型选择

2. DCT

帧内块AC系数采用与I帧相同方法的变换得到。帧内块DC系数变换方法有所不同，其预测值设为1024，除非先前的块是采用帧内编码。

非帧内块系数表示预测像素值之间的差异，而不是实际像素值。它们由当前宏块的值减去前一帧的运动补偿值得到。不预测DC值。

548

输入像素值为－255~＋255，输出结果为范围从－2000~＋2000的AC系数。

3. 量化

帧内块与I帧中的量化方法相同。

非帧内块采用量化器刻度和非帧内量化矩阵进行量化。AC和DC系数采用相同方式进行量化。

4. 帧内块编码

帧内块与I帧帧内块编码方式相同。在处理DC系数的时候有一点差异，其预测值为128，除非先前块采用帧内编码。

5. 非帧内块编码

编码块模式（Coded Block Pattern，CBP）用于指示哪些块有系数数据。除了DC系数与AC系数采用相同的编码方式以外，其编码方式与帧内块编码方式类似。

12.4.7　B帧

1. 宏块

根据后向运动补偿附加的复杂程度不同，B帧有12种宏块类型，如表12-11所示。

不编码宏块（skipped）为与前一个宏块有相同运动矢量和宏块类型的宏块，它不能采用帧内编码方式编码。不编码宏块的优点是它们几乎不占用传输带宽，不用编码，只需对宏块地址增量

进行编码，从而跳过对该宏块的编码。

如果表12-11中的[macroblock quant]项是1，将传输量化器刻度。而对于剩下的宏块类型，则采用前一个量化器刻度对DCT修正进行编码。

549

如果表12-11中的[motion forward]项是1，将要相继传输水平和垂直前向运动矢量。如果表12-11中的[motion backward]项是1，将要相继传输水平和垂直后向运动矢量。如果前向后向运动类型都出现了，运动矢量按照以下顺序传输：

水平前向→垂直前向→水平后向→垂直后向

如果表12-11中的[coded pattern]项是1，6位长的编码块模式将采用变长编码方式传输，并指示解码器宏块的6个块中哪一块已经被编码（设为1），哪一块还没有被编码。表12-14列出63种可能组合对应的码字。如果还没有块被编码，就没有码字。编码块模式由宏块类型决定。对于I帧中的宏块和P帧、B帧中的帧内编码宏块，编码块模式不需要传输，且假定为63（表示所有块都已经被编码）。

为确定采用何种类型的宏块，编码器需要做出一系列决策，如图12-4所示。

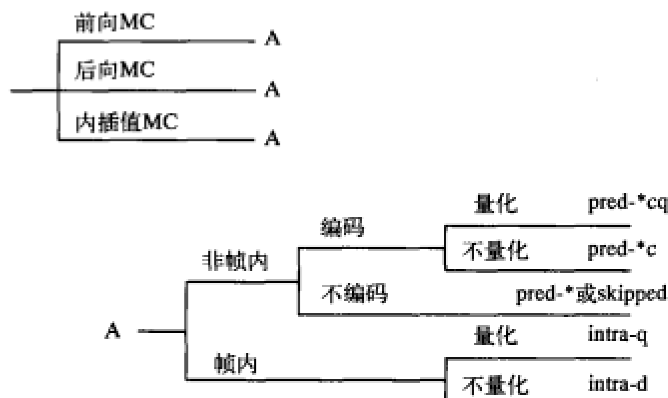


图12-4 MPEG-1 B帧宏块类型选择

2. 编码

块中的DCT系数将转换成量化系数，并且其编码方式与P帧的编码方式相同。

12.4.8 D帧

D帧只包含DC频率数据，一般用于快速可视化搜索应用。D帧包含的数据应当足以确保帮助用户定位所需的视频。

550

12.5 视频比特流

图12-5对视频比特流进行了举例说明，从上到下总共7个层次，各层分别是：

视频序列
序列头
图像组（GOP）
图像
条带
宏块（MB）
块

注意，起始码（0x000001xx）必须是字节对齐的，其实现方式是在高位之前插入0~7个0达到。

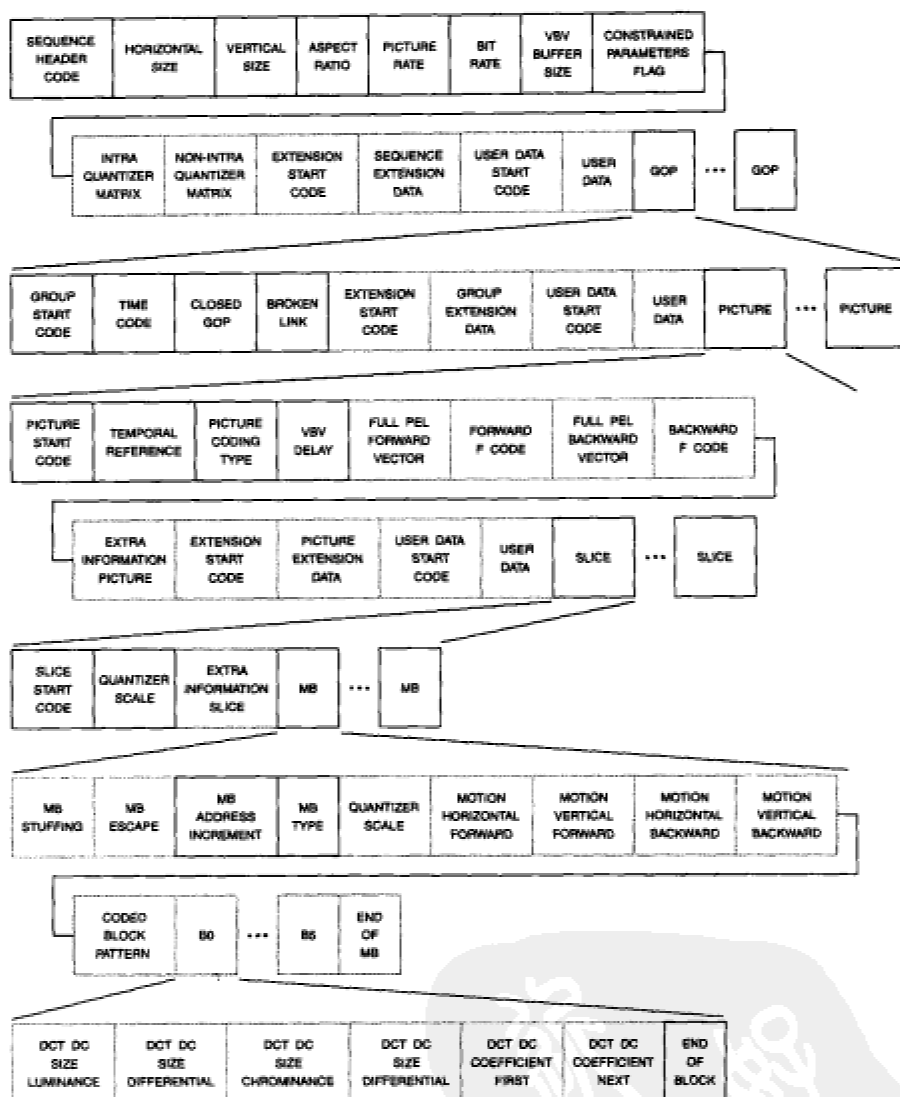


图12-5 MPEG-1视频比特流层次结构。标记位和保留位没有给出

12.5.1 视频序列

sequence_end_code

长度为32位的字段，值为0x000001B7，标志一个视频序列的结束。

12.5.2 序列头

每个序列的数据由序列头和随后的图像组（GOP）组成。结构如图12-5所示。

1. sequence_header_code

长度为32位的字段，值为0x000001B3，表示序列头的开始。

2. horizontal_size

长度为12位的二进制值，表示Y分量可视部分的宽度。该宽度在宏块中定义为(horizontal_size

+15)/16。

3. vertical_size

长度为12位的二进制值，表示Y分量可视部分的高度。该高度在宏块中定义为(vertical_size+15)/16。

4. pel_aspect_ratio

长度为4位的码字，表示像素的宽高比，如表12-3所示。

表12-3 MPEG-1 pel_aspect_ratio码字

高度/宽度	示　例	宽高比编码
禁止		0000
1.0000	方形像素	0001
0.6735		0010
0.7031	576行16：9	0011
0.7615		0100
0.8055		0101
0.8437	480行16：9	0110
0.8935		0111
0.9157	576行4：3	1000
0.9815		1001
1.0255		1010
1.0695		1011
1.0950	480行4：3	1100
1.1575		1101
1.2015		1110
保留		1111

5. picture_rate

长度为4位的码字，表示帧率，如表12-4所示。

表12-4 MPEG-1 picture_rate码字

帧/秒	图像速率编码	帧/秒	图像速率编码
禁止	0000	60	1000
24/1.001	0001	保留	1001
24	0010	保留	1010
25	0011	保留	1011
30/1.001	0100	保留	1100
30	0101	保留	1101
50	0110	保留	1110
60/1.001	0111	保留	1111

6. bit_rate

长度为18位的二进制值，表示比特流的比特率，以400bit/s为单位度量，不足时取上界。该值不允许为0。当值为0x3FFFF时，表示此时为可变比特率操作。如果constrained_parameters_flag为1，那么比特率必须小于等于1.856Mbit/s。

7. marker_bit

恒为1。

8. vbv_buffer_size

长度为10位的二进制数，表示解码器对视频解码时，校验缓冲需要的最小存储容量。按照如下定义：

551～553

$B = 16 \times 1024 \times vbv_buffer_size$

如果constrained_parameters_flag为1，vbv_buffer_size必须小于等于40KB。

9. constrained_parameters_flag

如果满足如下条件，此位设为1：

horizontal_size≤768像素点

vertical_size≤576行

[(horizontal_size+15)/16] × [(vertical_size+15)/16] ≤396

[(horizontal_size+15)/16] × [(vertical_size+15)/16] ×picture_rate≤396*25

picture_rate≤30帧/秒

forward_f_code≤4

backward_f_code≤4

10. load_intra_quantizer_matrix

如果其后紧接intra_quantizer_matrix，此位设为1。如果设为0，采用如下默认值，直到下一个序列头出现：

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

11. intra_quantizer_matrix

可选的64个8位值列表，用来替换当前的帧内量化器值。不允许为0。intra_quant[0,0]通常为8。这些数值一直有效，直到下一个序列头出现。

12. load_non_intra_quantizer_matrix

如果其后紧跟non_intra_quantizer_matrix，此位设为1。如果设为0，采用如下默认值，直到下一个序列头出现：

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

13. non_intra_quantizer_matrix

可选的64个8位值列表，用来替换当前的非帧内量化器值。不允许为0。这些数值一直有效，直到下一个序列头出现。

14. extension_start_code

可选的32位字符串，值为0x000001B5，表示sequence_extension_data（序列扩展数据）的开始。

sequence_extension_data一直持续，直到探测到下一个起始码。

15. sequence_extension_data

长度为$n \times 8$位，当extension_start_code出现时才出现。

16. user_data_start_code

可选的32位字符串，值为0x000001B2，表示user_data（用户数据）的开始，user_data一直持续，直到探测到下一个起始码。

17. user_data

长度为有$n \times 8$位，当user_data_start_code出现时才出现。user_data不能包含超过连续23个0的字符串。

554

12.5.3 图像组层

每个图像组由图像组头及后面跟着的图像数据组成。结构如图12-5所示。

1. group_start_code

长度为32位，值为0x000001B8，表示图像组的开始。

2. time_code

长度为25位，表示时间编码信息，如表12-5所示。当且仅当图像帧率为30/1.001（29.97）Hz时，[drop_frame_flag]设为1。

表12-5 MPEG-1 time_code字段

时间编码	取值范围	位数
drop_frame_flag		1
time_code_hours	0~23	5
time_code_minutes	0~59	6
market_bit	1	1
time_code_seconds	0~59	6
time_code_pictures	0~59	6

3. closed_gop

1位标志位，如果图像组没有参考前一个的图像组运动矢量而进行编码，此位设为1。此位允许支持编辑压缩比特流。

4. broken_link

1位标志位，编码时此位设为0。在编辑过程中，当图像组第一个I帧后面的B帧不能正确地解码时，此位设为1。

5. extension_start_code

可选的长度为32位的字符串，值为0x000001B5，表示group_extension_data（组扩展数据）。group_extension_data一直持续，直到下一个起始码出现。

6. group_extension_data

当extension_start_code出现时才出现，长度为$n \times 8$位。

7. user_data_start_code

可选的长度为32位的字符串，值为0x000001B2，表示user_data的开始，user_data一直持续，直到下一个起始码出现。

8. user_data

当user_data_start_code出现时才出现，长度为$n \times 8$位。user_data不能包含23个或更多个连续0的字符串。

555

12.5.4 图像层

每个图像层的数据由图像头和紧跟其后的条带数据组成。其结构如图12-5所示。

1. picture_start_code

长度为32位，值为0x00000100，表示图像的开始。

2. temporal_reference

长度为10位的二进制，对于每个图像组显示顺序中的第一帧，temporal_reference值为0。对于显示顺序中的其他每一帧，该值以1024为模，每次递增1。

3. picture_coding_type

长度为3位的码字，表示帧类型（I帧、P帧、B帧或D帧），如表12-6所示。D帧不与其他帧在同一视频序列使用。

4. vbv_delay

长度为16位的二进制值。对于恒定比特率，vbv_delay设定了解码图像开始时候缓存的占用空间，以免出现溢出或者下溢情况。对可变比特率，vbv_delay值恒为0xFFFF。

5. full_per_forward_vector

1位标志位，当picture_coding_type为010（P帧）或011（B帧）时，该位出现。若值为1，则前向运动矢量基于整像素，而不是半像素：

6. forward_f_code

长度为3位的码字，在picture_coding_type为010（P帧）或011（B帧）时出现，取值范围为001到111，不允许取值为000。

解码器采用这个字段派生出的两个参数解码前向运动矢量：forward_r_size和forward_f。forward_r_size比forward_f_code小1，表12-7中给出了forward_f的定义。

表12-6 MPEG-1 picture_coding_type编码

编码类型	编码
禁用	000
I帧	001
P帧	010
B帧	011
D帧	100
保留	101
保留	110
保留	111

表12-7 MPEG-1 forward_f_code值

forward_f_code	forward_f_code值
001	1
010	2
011	4
100	8
101	16
110	32
111	64

556

7. full_pel_backward_vector

1位标志位，在picture_coding_type为011（B帧）时出现。值为1时，后向运动矢量基于整像素，而不是半像素。

8. backward_f_code

长度为3位的二进制数，在picture_coding_type为011（B帧）时出现。取值范围为001~111，000禁止使用。

解码器采用这个字段派生出的两个参数解码后向运动矢量：backward_r_size和backward_f。backward_r_size比backward_f_code小1，backward_f定义与forward_f相同。

9. extra_bit_picture

1位标志位，设为1时，表示后面跟有extra_information_picture（附加信息图像）。

10. extra_information_picture

如果extra_bit_picture设为1，这9位将出现其后，其中包括8位数据（extra_information_picture）和1位extra_bit_picture，表明是否还跟着一个9位的extra_information_picture，如此继续。

11. extension_start_code

可选的长度为32位的字符串，值为0x000001B5，表示picture_extension_data（图像扩展数据）的开始，picture_extension_data一直持续，直到下一个起始码出现。

12. picture_extension_data

在extension_start_code出现时才出现，长度为$n \times 8$位。

13. user_data_start_code

可选的长度为32位的字符串，值为0x000001B2，表示user_data的开始，user_data一直持续，直到下一个起始码出现。

14. user_data

在user_data_start_code出现时才出现，长度为$n \times 8$位。用户数据不能包含23个以上连续0的字符串。

12.5.5 条带层

每个条带层数据由条带头和紧跟其后的宏块数据组成。其结构如图12-5所示。

1. slice_start_code

长度为32位的字段，前24位值为0x000001。后8位为slice_vertical_position，值为0x01~0xAF。

slice_vertical_position表示条带中第一个宏块在宏块组（macroblock unit）中的垂直位置。宏块的第一行值为1。

2. quantizer_scale

长度为5位的二进制数，取值范围为1~31（不允许为0）。指定DCT系数重构非零值的比例因子。解码器一直使用这个值，直到在条带层或者宏块层中接收到另一个quantizer_scale。

3. extra_bit_slice

1位标志位，设为1时表示后面跟着extra_information_slice（附加信息条带）。 557

4. extra_information_slice

如果extra_bit_slice为1，这9位将出现其后，其中包括8位的数据（extra_information_slice）和1位的extra_bit_slice，表明是否在其后还跟着另一个9位的extra_information_slice。如此反复。

12.5.6 宏块层

每个宏块层的数据由宏块头和紧跟其后的运动矢量和块数据组成。其结构如图12-5所示。

1. macroblock_stuffing

可选的长度为11位的固定比特字符串，值为0000 0001 111，可用于提高比特率以满足存储和传输的需要。可以采用任意个数的连续macroblock_stuffing。

2. macroblock_escape

可选的长度为11位的固定比特字符串，值为0000 0001 000，如果当前宏块地址和前一个宏块地址的差值超过33，将采用此项。它强制macroblock_adress_increment的值变为33。可以采用任意个数的连续macroblock_escape。

3. macroblock_address_increment

可变长码字，指明当前宏块和前一个宏块的地址差值。最大值为33。超过33的值采用macroblock_escape编码。表12-8列出了这些可变长编码。

表12-8 MPEG-1 macroblock_address_increment可变长编码表

增 量 值	编 码	增 量 值	编 码
1	1	17	0000 0101 10
2	011	18	0000 0101 01
3	010	19	0000 0101 00
4	0011	20	0000 0100 11
5	0010	21	0000 0100 10
6	0001 1	22	0000 0100 011
7	0001 0	23	0000 0100 010
8	0000 111	24	0000 0100 001
9	0000 110	25	0000 0100 000
10	0000 1011	26	0000 0011 111
11	0000 1010	27	0000 0011 110
12	0000 1001	28	0000 0011 101
13	0000 1000	29	0000 0011 100
14	0000 0111	30	0000 0011 011
15	0000 0110	31	0000 0011 010
16	0000 0101 11	32	0000 0011 001
		33	0000 0011 000

4. macroblock_type

可变长码字，指明编码方法和宏块类型。如表12-9~表12-12所示。

表12-9 MPEG-1 I帧macroblock_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
intra-d	0	0	0	0	1	1
intra-q	1	0	0	0	1	01

表12-10 MPEG-1 P帧macroblock_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-mc	0	1	0	1	0	1
pred-c	0	0	0	1	0	01
pred-m	0	1	0	0	0	001
intra-d	0	0	0	0	1	0001 1
pred-mcq	1	1	0	1	0	0001 0
pred-cq	1	0	0	1	0	0000 1
intra-q	1	0	0	0	1	0000 01
skipped						

表12-11 MPEG-1 B帧macroblock_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-i	0	1	1	0	0	10
pred-ic	0	1	1	1	0	11
pred-b	0	0	1	0	0	010
intra-bc	0	0	1	1	0	011

（续）

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-f	0	1	0	0	0	0010
pred-fc	0	1	0	1	0	0011
intra-d	0	0	0	0	1	0001 1
pred-icq	1	1	1	1	0	0001 0
pred-fcq	1	1	0	1	0	0000 11
pred-bcq	1	0	1	1	0	0000 10
intra-q	1	0	0	0	1	0000 01
skipped						

表12-12 MPEG-1 D帧macroblock_type可变长编码表

Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编 码
0	0	0	0	1	1

5. quantizer_scale

可选的长度为5位的二进制数，值为1~31（不允许为0），指示接收的DCT系数重构非零值的比例因子。解码器在接收下一个来自条带层或宏块层的quantizer_scale前始终采用当前值。quantizer_scale仅在[macroblock quant]为1时出现，如表12-9~表12-12所示。

6. motion_horizontal_forward_code

可选的变长码字，包含前向运动矢量信息，如表12-13所示。当[motion forward]为1时出现，如表12-9~表12-12所示。

558

7. motion_horizontal_forward_r

可选的二进制数（属于forward_r_size位），用于帮助解码前向运动矢量。当[motion forward]为1时出现，如表12-9~表12-12所示，forward_f_code≠001且motion_horizontal_forward_code≠0。

8. motion_vertical_forward_code

可选的变长码字，包含如表12-13所示的前向运动矢量信息，当[motion forward]为1时出现，如表12-9~表12-12所示。

表12-13 MPEG-1 motion_horizontal_forward_code、motion_vertical_forward_code、motion_horizontal_backward_code和motion_vertical_backward_code可变长编码表

运动矢量差分	编 码	运动矢量差分	编 码
−16	0000 0011 001	1	010
−15	0000 0011 011	2	0010
−14	0000 0011 101	3	000 10
−13	0000 0011 111	4	0000 110
−12	0000 0100 001	5	0000 1010
−11	0000 0100 011	6	0000 1000
−10	0000 0100 11	7	0000 0110
−9	0000 0101 01	8	0000 0101 10
−8	0000 0101 11	9	0000 0101 00
−7	0000 0111	10	0000 0100 10
−6	0000 1001	11	0000 0100 010
−5	0000 1011	12	0000 0100 000

（续）

运动矢量差分	编　码	运动矢量差分	编　码
−4	0000 111	13	0000 0011 110
−3	0001 1	14	0000 0011 100
−2	0011	15	0000 0011 010
−1	011	16	0000 0011 000
0	1		

9. motion_vertical_forward_r

可选的二进制数（属于forward_r_size位），用于帮助解码前向运动矢量。当表12-9~表12-12中的[motion forward]为1时出现。forward_f_code≠001且motion_vertical_forward_code≠0。

10. motion_horizontal_backward_code

可选的变长码字，包含表12-13中定义的后向运动矢量信息。当表12-9~表12-12中的[motion backward]为1时出现。

11. motion_horizontal_backward_r

可选的二进制数（属于backward_r_size位），用于帮助解码后向运动矢量。当表12-9~表12-12中的[motion backward]为1时出现。backward_f_code≠001且motion_horizontal_backward_code≠0。

12. motion_vertical_backward_code

可选的变长码字，包含如表12-13中定义的后向运动矢量信息，解码值帮助决定motion_vertical_backward_r是否在比特流出现。当表12-9~表12-12中的[motion backward]为1时出现。

13. motion_vertical_backward_r

可选的二进制数（属于backward_r_size位），用于帮助解码后向运动矢量。当表12-9~表12-12中的[motion backward]为1时出现。backward_f_code≠001且motion_vertical_backward_code≠0。

14. coded_block_pattern

可选的变长码字，用于表示如表12-14所示的编码块模式。当表12-9~表12-12中的[coded pattern]为1时候出现。指示宏块中哪些（至少一个）变换系数需要传输。编码块模式的二进制编码采用如下形式表示：

$P_0P_1P_2P_3P_4P_5$

当块第[*n*]个系数出现时，对应的P_n为1，否则为0。图12-2给出块编号（十进制格式）方法。

表12-14 MPEG-1 coded_block_pattern可变长编码表

编码块模式	编　码	编码块模式	编　码	编码块模式	编　码
60	111	9	0010 110	43	0001 0000
4	1101	17	0010 101	25	0000 1111
8	1100	33	0010 100	37	0000 1110
16	1011	6	0010 011	26	0000 1101
32	1010	10	0010 010	38	0000 1100
12	1001 1	18	0010 001	29	0000 1011
48	1001 0	34	0010 000	45	0000 1010
20	1000 1	7	0001 1111	53	0000 1001
40	1000 0	11	0001 1110	57	0000 1000
28	0111 1	19	0001 1101	30	0000 0111
44	0111 0	35	0001 1100	46	0000 0110

（续）

编码块模式	编　码	编码块模式	编　码	编码块模式	编　码
52	0110 1	13	0001 1011	54	0000 0101
56	0110 0	49	0001 1010	58	0000 0100
1	0101 1	21	0001 1001	31	0000 0011 1
61	0101 0	41	0001 1000	47	0000 0011 0
2	0100 1	14	0001 0111	55	0000 0010 1
62	0100 0	50	0001 0110	59	0000 0010 0
24	0011 11	22	0001 0101	27	0000 0001 1
36	0011 10	42	0001 0100	39	0000 0001 0
3	0011 01	15	0001 0011		
63	0011 00	51	0001 0010		
5	0010 111	23	0001 0001		

15. end_of_macroblock

可选的1位标志位，值为1。仅在D帧中出现。

12.5.7 块层

每个块层数据都由系数数据组成，结构如图12-5所示。

1. dct_dc_size_luminance

可选的可变长码字，用于Y块帧内编码。指示用于dct_dc_differential的位数目。表12-15列出了这些变长码字。

559 ~ 562

表12-15 MPEG-1 dct_dc_size_luminance可变长编码表

dct_dc_size_luminance	编　码	dct_dc_size_luminance	编　码
0	100	5	1110
1	00	6	1111 0
2	01	7	1111 10
3	101	8	1111 110
4	110		

563

2. dct_dc_differential

可选的可变长码字，当dct_dc_size_luminance不为0时，在dct_dc_size_luminance之后出现，其值如表12-16所示。

表12-16 MPEG-1 dct_dc_differential可变长编码表

dct_dc_differential	Size	编码（Y）	编码（CbCr）	附加编码
−255~−128	8	1111110	11111110	00000000~01111111
−127~−64	7	111110	1111110	0000000~0111111
−63~−32	6	11110	111110	000000~011111
−31~−16	5	1110	11110	00000~01111
−15~−8	4	110	1110	0000~0111
−7~−4	3	101	110	000~011
−3~−2	2	01	10	00~01
−1	1	00	01	0

（续）

dct_dc_differential	Size	编码（Y）	编码（CbCr）	附加编码
0	0	100	00	
1	1	00	01	1
2~3	2	01	10	10~11
4~7	3	101	110	100~111
8~15	4	110	1110	1000~1111
16~31	5	1110	11110	10000~11111
32~63	6	11110	111110	100000~111111
64~127	7	111110	1111110	1000000~1111111
128~255	8	1111110	11111110	10000000~11111111

3. dct_dc_size_chrominance

可选的变长码字，用于Cb、Cr块的帧内编码。指示用于dct_dc_differential的位数目。表12-17列出了这些可变长码字。

表12-17 MPEG-1 dct_dc_size_chrominance可变长编码表

dct_dc_size_chrominance	编 码	dct_dc_size_chrominance	编 码
0	00	5	1111 0
1	01	6	1111 10
2	10	7	1111 110
3	110	8	1111 1110
4	1110		

4. dct_dc_differential

可选的可变长码字，当dct_dc_size_chrominance不为0时，在dct_dc_size_chrominance之后出现。其值如表12-16所示。

5. dct_coefficient_first

可选的可变长码字，用于非帧内编码块的第一个DCT系数，其定义在表12-18和表12-19中给出。

表12-18 MPEG-1 dct_coefficient_first和dct_coefficient_next可变长编码表

RUN	LEVEL	编 码[1]	RUN	LEVEL	编 码
end_of_block		10	escape		0000 01
0[2]	1	1 s	0	5	0010 0110 s
0[3]	1	11 s	0	6	0010 0001 s
1	1	011 s	1	3	0010 0101 s
0	2	0100 s	3	2	0010 0100 s
2	1	0101 s	10	1	0010 0111 s
0	3	0010 1 s	11	1	0010 0011 s
3	1	0011 1 s	12	1	0010 0010 s
4	1	0011 0 s	13	1	0010 0000 s
1	2	0001 10 s	0	7	0000 0010 10 s
5	1	0001 11 s	1	4	0000 0011 00 s
6	1	0001 01 s	2	3	0000 0010 11 s
7	1	0001 00 s	4	2	0000 0011 11 s
0	4	0000 110 s	5	2	0000 0010 01 s

（续）

RUN	LEVEL	编　码[1]	RUN	LEVEL	编　码
end_of_block		10	escape		0000 01
2	2	0000 100 s	14	1	0000 0011 10 s
8	1	0000 111 s	15	1	0000 0011 01 s
9	1	0000 101 s	16	1	0000 0010 00 s
0	8	0000 0001 1101 s	0	12	0000 0000 1101 0 s
0	9	0000 0001 1000 s	0	13	0000 0000 1100 1 s
0	10	0000 0001 0011 s	0	14	0000 0000 1100 0 s
0	11	0000 0001 0000 s	0	15	0000 0000 1011 1 s
1	5	0000 0001 1011 s	1	6	0000 0000 1101 0 s
2	4	0000 0001 0100 s	1	7	0000 0000 1010 1 s
3	3	0000 0001 1100 s	2	5	0000 0000 1010 0 s
4	3	0000 0001 0010 s	3	4	0000 0000 1001 1 s
6	2	0000 0001 1110 s	5	3	0000 0000 1001 0 s
7	2	0000 0001 0101 s	9	2	0000 0000 1000 1 s
8	2	0000 0001 0001 s	10	2	0000 0000 1000 0 s
17	1	0000 0001 1111 s	22	1	0000 0000 1111 1 s
18	1	0000 0001 1010 s	23	1	0000 0000 1111 0 s
19	1	0000 0001 1001 s	24	1	0000 0000 1110 1 s
20	1	0000 0001 0111 s	25	1	0000 0000 1110 0 s
21	1	0000 0001 0110 s	26	1	0000 0000 1101 1 s
0	16	0000 0000 0111 11 s	0	40	0000 0000 0010 000 s
0	17	0000 0000 0111 10 s	1	8	0000 0000 0011 111 s
0	18	0000 0000 0111 01 s	1	9	0000 0000 0011 110 s
0	19	0000 0000 0111 00 s	1	10	0000 0000 0011 101 s
0	20	0000 0000 0110 11 s	1	11	0000 0000 0011 100 s
0	21	0000 0000 0110 10 s	1	12	0000 0000 0011 011 s
0	22	0000 0000 0110 01 s	1	13	0000 0000 0011 010 s
0	23	0000 0000 0110 00 s	1	14	0000 0000 0011 001 s
0	24	0000 0000 0101 11 s	1	15	0000 0000 0001 0011 s
0	25	0000 0000 0101 10 s	1	16	0000 0000 0001 0010 s
0	26	0000 0000 0101 01 s	1	17	0000 0000 0001 0001 s
0	27	0000 0000 0101 00 s	1	18	0000 0000 0001 0000 s
0	28	0000 0000 0100 11 s	6	3	0000 0000 0001 0100 s
0	29	0000 0000 0100 10 s	11	2	0000 0000 0001 1010 s
0	30	0000 0000 0100 01 s	12	2	0000 0000 0001 1001 s
0	31	0000 0000 0100 00 s	13	2	0000 0000 0001 1000 s
0	32	0000 0000 0011 000 s	14	2	0000 0000 0001 0111 s
0	33	0000 0000 0011 111 s	15	2	0000 0000 0001 0110 s
0	34	0000 0000 0011 110 s	16	2	0000 0000 0001 0101 s
0	35	0000 0000 0010 101 s	27	1	0000 0000 0001 1111 s
0	36	0000 0000 0010 100 s	28	1	0000 0000 0001 1110 s
0	37	0000 0000 0010 011 s	29	1	0000 0000 0001 1101 s
0	38	0000 0000 0010 010 s	30	1	0000 0000 0001 1100 s
0	39	0000 0000 0010 001 s	31	1	0000 0000 0001 1011 s

注：1. s表示非零系数（level）的符号，0表示正，1表示负。

2. 用于dct_coefficient_first。

3. 用于dct_coefficient_next。

表12-19　dct_coeffient_first和dct_coefficient_next中转义码后的游程、非零系数编码

RUN	LEVEL	固定长度编码
0		0000 00
1		0000 01
2		0000 10
⋮		⋮
63		1111 11
	−256	禁止
	−255	1000 0000 0000 0001
	−254	1000 0000 0000 0010
	⋮	⋮
	−129	1000 0000 0111 1111
	−128	1000 0000 1000 0000
	−127	1000 0001
	−126	1000 0010
	⋮	⋮
	−2	1111 1110
	−1	1111 1111
	0	禁止
	1	0000 0001
	⋮	⋮
	127	0111 1111
	128	0000 0000 1000 0000
	129	0000 0000 1000 0001
	⋮	⋮
	255	0000 0000 1111 1111

6. dct_coefficient_next

最大为63的可选的可变长码字，仅用于I帧、P帧和B帧。表示在第一个系数之后的DCT系数，其定义在表12-18和表12-19中给出。

7. end_of_block

564 长度为2位，仅在I帧、P帧和B帧中出现，表示其后面没有其他非零系数。该参数值为10。

12.6　系统比特流

系统比特流将视频音频比特流复用到一个单独的比特流，并且与控制信息一起格式化成MPEG-1定义的指定协议。

包数据可能包含视频或音频信息。至多可复用16路视频和32路音频，并支持两种类型的私有数据流。一种是完全私有的，另一种用于支持同步和缓冲管理。

尽管数据包长度可以非常大，但通常最大的长度为2048字节。在CD-ROM存储时，数据包长度与扇区大小一致。典型情况是6个或7个视频数据包附带一个音频数据包。

图12-6举例说明了系统比特流的3层等级结构，自顶向下分别是：

- ISO/IEC 11172层
- 包组（Pack）
- 包（Packet）

注意，起始码（0x000001xx）必须为字节对齐，实现方式是在起始码之前插入0~7个0位。

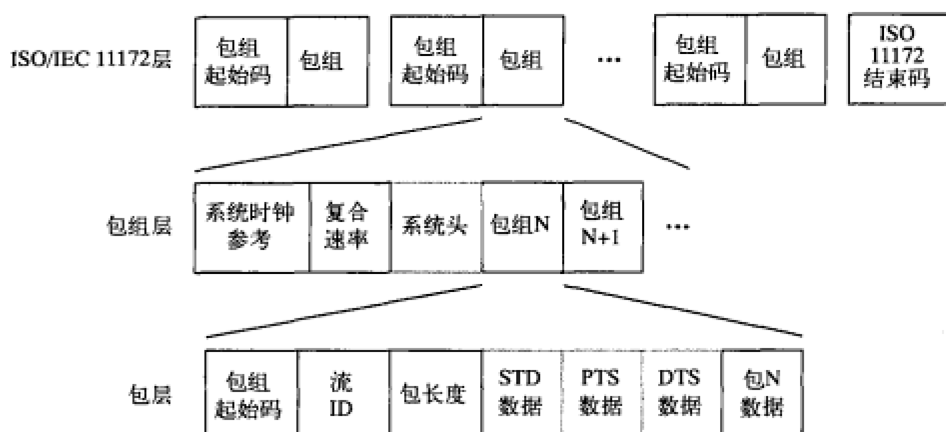


图12-6 MPEG-1系统比特流层结构，标记位和保留位没有给出

12.6.1 ISO/IEC 11172层

ISO_11172_end_code

长度为32位的字段，值为0x000001B9，用于终止一段系统比特流。

12.6.2 包组层

每个包组由包组头和紧跟其后的一个系统头（可选的）和包数据组成。结构如图12-6所示。

1. pack_start_code

长度为32位的字段，值为0x000001BA，表示包组的开始。

2. fixed_bits

长度为4位的二进值数，值恒为0010。

3. system_clock_reference_32~30

system_clock_reference（SCR）为长度为33位的数，该数由标记位分成的三部分编码而成。

system_clock_reference表示system_clock_reference字段的最后一个字节到达解码器输入处的预计时间。system_clock_reference的值为90kHz时钟周期倍数。

4. marker_bit

标记位，值恒为1。

5. system_clock_reference_29~15

6. marker_bit

标记位，值恒为1。

7. system_clock_reference_14~0

8. marker_bit

标记位，值恒为1。

9. marker_bit

标记位，值恒为1。

565 ~ 570

10. mux_rate

长度为22位的二进制数，指示解码器接收比特流的速率。以50字节/秒为单位，不足的取上

界。不允许为0。

11. marker_bit

标记位，值恒为1。

12.6.3 系统头

1. system_header_start_code

长度为32位的字段，值为0x000001BB，表示系统头的开始。

2. header_length

长度为16位的二进制数，指示跟在系统头之后的字节数。

3. marker_bit

标记位，值恒为1。

4. rate_bound

长度为22位的二进制数，指定一个大于等于mux_rate最大值的整数值，解码器可以据此确定是否有能力对整个比特流解码。

5. marker_bit

标记位，值恒为1。

6. audio_bound

长度为6位的二进制数，取值范围为0~32，指定一个大于等于同时活动的音频流最大数的整数值。

7. fixed_flag

571

1位标志位，指示操作是固定比特率（为1）还是可变比特率（为0）。

8. CSPS_flag

1位标志位，指示比特流是约束系统参数流（为1），否则为0。

9. system_audio_lock_flag

如果音频采样率和解码器系统时钟频率保持固定关系，该位值为1。

10. system_video_lock_flag

如果视频图像速率和解码器系统时钟频率保持固定关系，该位值为1。

11. marker_bit

标记位，值恒为1。

12. video_bound

长度为5位的二进制数，取值范围为0~16，指定一个大于等于同时活动的视频流最大数目的整数值。

13. reserved_byte

长度为8位，通常为1111 1111。

14. stream_ID

可选的8位的字段，如表12-20定义。表示接下来的与STD_buffer_bound_scal和STD_buffer_size_bound相关的类型和流编号。每个系统比特流中出现的视频和音频流在系统头中只出现一次。

表12-20 MPEG-1 stream_ID码

流　类　型	流ID
所有音频流	1011 1000
所有视频流	1011 1001
保留流	1011 1100
私有流1	1011 1101
填充流	1011 1110
私有流2	1011 1111
音频流编号xxxxx	110x xxxx
视频流编号xxxx	1110 xxxx
保留数据流编号xxxx	1111 xxxx

15. fixed_bits

可选的长度为2位的字段，值为11。当stream_ID出现时才出现。

16. STD_buffer_bound_scale

可选的1位字段，指示用于解释STD_buffer_size_bound的缩放因子。对于音频流，该字段值为0；对于视频流，该字段值为1。对其他流类型，1和0都可以。当stream_ID出现时该字段才出现。 572

17. STD_buffer_size_bound

可选的长度为13位的二进制数，指定一个大于或等于解码器输入缓冲最大值的整数值。当STD_buffer_bound_scale为0时，STD_buffer_size_bound以128字节为单位。当STD_buffer_bound_scale为1时，以1024字节为单位。该字段仅当stream_ID出现时才出现。

12.6.4 包层

1. packet_start_code_prefix

长度为24位的字段，值为0x000001，跟在stream_ID相应的参数项后面，表示包的开始。

2. stream_ID

长度为8位的二进制数，指定当前比特流的类型和数量。如表12-20所示。

3. packet_length

长度为16位二进制数，指定包的字节数。

4. stuffing_byte

可选参数，值为1111 1111。根据存储媒介需要，至多可使用16个连续的stuffing_bytes。当stream_ID不为私有流2时出现。

5. STD_bits

可选的长度为2位二进制数，值为01，表示接下来跟着STD_buffer_scale和STD_buffer_size。当stream_ID不为私有流2时出现.

6. STD_buffer_scale

可选的1位字段，指示解释STD_buffer_size的比例因子，值为0时表示音频流，为1时表示视频流，对其他类型流，0和1都可以。当stream_ID不为私有流2时出现。

7. STD_buffer_size

可选的长度为13位的二进制数，指定解码器输入缓冲的大小。当STD_buffer_scale为0时，STD_buffer_size以128字节为单位，当STD_buffer_scale为1，以1024字节为单位。当stream_ID不为私有流2时出现。

8. PTS_bits

可选的4位二进制数，值为0010，表示接下来将出现展示时间戳选项，当stream_ID不为私有流2时出现。

9. presentation_time_stamp_32~30

可选的33位二进制数，presentation_time_stamp（PTS）由标记位分成3段来编码，表示解码器显示的计划时间。PTS取值为90kHz系统时钟的倍数。当PTS_bits出现且stream_ID不为私有流2时出现。

10. marker_bit

可选标记位，值恒为1，当PTS_bits出现且stream_ID不为私有流2的时候出现。

11. presentation_time_stamp_29~15

可选，当PTS_bits出现且stream_ID不为私有流2的时候出现。

12. marker_bit

573 可选的标记位，值恒为1。当PTS_bits出现且stream_ID不为私有流2时出现。

13. presentation_time_stamp_14~0

可选字段，当PTS_bits出现且stream_ID不为私有流2的时候出现。

14. marker_bit

可选标记位，值恒为1。当PTS_bits出现且stream_ID不为私有流2时出现。

15. DTS_bits

可选的长度为4位的二进制数，值为0011，表示接下来将出现展示和解码时间戳选项。当stream_ID不为私有流2时出现。

16. presentation_time_stamp_32~30

可选的长度为33位数，presentation_time_stamp（PTS）由标记位分成3段来编码，表示解码显示的计划时间。PTS取值为90kHz系统时钟的倍数。当DTS_bits出现且stream_ID不为私有流2时出现。

17. marker_bit

可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

18. presentation_time_stamp_29~15

可选字段，当DTS_bits出现且stream_ID不为私有流2时出现。

19. marker_bit

可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

20. presentation_time_stamp_14~0

可选字段，当DTS_bits出现且stream_ID不为私有流2时出现。

21. marker_bit

可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

22. fixed_bits

可选长度为4位的二进制数，值为0001，当DTS_bits出现且stream_ID不为私有流2时出现。

23. decoding_time_stamp_32~30

可选的长度为33位的二进制数，decoding_time_stamp（DTS）由标记位分成3段来编码，表示解码显示的计划时间。DTS取值为90kHz系统时钟的倍数。当DTS_bits出现且stream_ID不为私有流2时出现。

24. marker_bit

可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

25. decoding_time_stamp_29~15

可选字段，当DTS_bits出现且stream_ID不为私有流2时出现。

26. marker_bit

574 可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

27. decoding_time_stamp_14~0

可选字段，当DTS_bits出现且stream_ID不为私有流2时出现。

28. marker_bit

可选标记位，值恒为1。当DTS_bits出现且stream_ID不为私有流2时出现。

29. nonPTS_nonDTS_bits

可选的长度为8位的二进制数，值为00001111，当STD_bits、PTS_bits或DTS_bits（及其相应

后续字段）不出现时出现。

30. packet_data_byte

由包层中的stream_ID指定的比特流中n字节数据，字节数由参数packet_length决定。

12.7 视频解码

系统解复用器分析系统比特流，解复用视频音频比特流。

视频解码器本质上执行与编码器相反的工作过程。它从编码视频比特流中重构I帧。利用I帧和附加编码数据、运动矢量等构建P帧和B帧。最后，把各帧按照合适的顺序输出。

12.7.1 快进播放注意事项

快进播放操作可采用D帧或仅对I帧解码来实现。然而，在高速率下仅对I帧解码会给传输媒质和解码器带来很大的负担。

另一种方法是，发送源可以仅挑选快放所需的I帧传输，使得比特率仍然保持稳定。

12.7.2 暂停模式注意事项

解码器必须能够控制输入的比特流，否则，在播放恢复的时候会出现延迟或跳帧现象。

12.7.3 快退播放注意事项

解码器必须能够在视频播放方向解码每一个图片组，并存储这些解码的图片组，这样才能实现快退播放。为最小化解码器的存储空间，图片组必须比较小，或者可以对帧进行重排序。重排序可以采用另一个顺序传输帧来实现，也可以通过在解码器缓存里重新调整图片组的顺序来实现。

12.7.4 解码后期处理

SIF数据通常转换为720×480i或720×576i。MPEG-1规范中考虑过隐式上采样过滤。原始解码线与场1对应。场2采用内插线。 575

12.8 现实问题

12.8.1 系统比特流终止

系统比特流中的sequence_end_code位置错误是一个常见的问题。如果这种情况发生了，一些解码器将不知道视频结束的位置，而输出乱七八槽的东西。

消除拖动帧使得系统比特流缩短，同时会一并去掉sequence_end_code关键字，此时会产生另外一个问题：解码器无法确定在什么时候停止播放。

12.8.2 时间码

由于一些解码器依赖时间码信息，所以必须有时间码措施。为让问题最小化，视频流开始时将时间码设为0，每一帧时间码加1。

12.8.3 可变比特率

虽然MPEG-1支持可变比特率，但如果可能还是尽量采用恒定比特率。由于vbv_delay对可变

比特率没意义，对可变比特率情况，MPEG-1标准指定其设为最大值。

然而，某些解码器针对可变比特率也使用vbv_delay参数，这样导致起始视频延迟2~3秒，头60~90帧数据都被跳过去了。

12.8.4 约束比特流

多数MPEG-1解码器只能处理一部分MPEG-1约束参数。为保证效果最好，只允许采用那一部分约束参数。

12.8.5 源采样时钟

要得到高压缩、低失真的效果，需要视频源产生或使用非常稳定的采样时钟。采用稳定时钟保证整幅图像样本的垂直对齐。质量差的时钟使得图像的失真更加剧烈。

参考文献

1. Digital Video Magazine, "*Not All MPEGs Are Created Equal*", by John Toebes, Doug Walker, and Paul Kaiser, August 1995.
2. Digital Video Magazine, "*Squeeze the Most From MPEG*", by Mark Magel, August 1995.
3. ISO/IEC 11172-1, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 1: Systems.*
4. ISO/IEC 11172-2, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 2: Video.*
5. ISO/IEC 11172-3, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 3: Audio.*
6. ISO/IEC 11172-4, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 4: Compliance testing.*
7. ISO/IEC 11172-5, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 5: Software simulation.*
8. Watkinson, John, *The Engineer's Guide to Compression*, Snell and Wilcox Handbook Series.

576



第13章　MPEG-2

MPEG-2扩展了MPEG-1的相关功能和特性，从而覆盖了更加广泛的应用范围。在阅读本章之前，请重新回顾一下第12章关于MPEG-1的有关内容，以熟悉MPEG的相关基础知识。

MPEG-2制定过程中所确定的主要目标应用是在4~9Mbit/s比特率下的全数字广播级视频传输。然而，MPEG-2对于许多其他应用也是非常有用的，如HDTV，并且MPEG-2现在可支持的比特率在1.5~60Mbit/s之间。

MPEG-2是一个ISO标准（ISO/IEC 13818），它由下列11部分组成：

系统	ISO/IEC 13818-1
视频	ISO/IEC 13818-2
音频	ISO/IEC 13818-3
一致性测试	ISO/IEC 13818-4
软件仿真	ISO/IEC 13818-5
DSM-CC扩展	ISO/IEC 13818-6
高级音频编码	ISO/IEC 13818-7
系统解码实时接口	ISO/IEC 13818-8
RTI扩展	ISO/IEC 13818-9
DSM-CC一致性	ISO/IEC 13818-10
IPMP	ISO/IEC 13818-11

与MPEG-1一样，MPEG-2中的压缩比特流隐式地定义了解压缩算法。而压缩算法则由各厂商独立开发，并允许在国际标准范围内获得所属权益。

数字存储媒体命令与控制（Digital Storage Media Command and Control，DSM-CC）扩展（ISO/IEC 13818-6）是一个用于开发与MPEG-2流相关的控制信道的工具集。除了提供VCR风格特征（如快进、快退、暂停等），DSM-CC还可用于各种其他目的的应用，如分组数据传输。DSM-CC将工作在下一代分组网络上，并与相关的因特网协议如RSVP、RSTP、RTP和SCP等一起工作。

实时接口（Real Time Interface，RTI）扩展（ISO/IEC 13818-9）定义了一个终端设备制造商和网络运营商都能设计的公共接口点。RTI在解码器的输入处为MPEG-2系统流的字节定义了一个传输模型，尽管MPEG-2系统定义了一个理想的字节传输时间表。

知识产权管理和保护（Intellectual Property Management and Protection，IPMP）是一个数字版权管理（Digital Right Management，DRM）标准，是根据MPEG-4的IPMP扩展规范改编过来的。IPMP并不是一个完整的系统，而是一个提供各种函数的框架。

13.1　音频概述

除了非向后兼容（non-backwards-compatible）音频扩展（ISO/IEC 13818-7）之外，MPEG-2还最多可支持5个与MPEG-1音频编码兼容的全带宽信道。MPEG-2同样也将MPEG-1的音频编码扩展到了半采样率（16kHz、22.05kHz和24kHz），从而改善在比特率小于等于每信道64kHz情况下的音频质量。

MPEG-2.5是一个对MPEG-2音频能力非官方的通用扩展，它增加了8kHz、11.025kHz和12kHz的采样率。

13.2　视频概述

在MPEG-2中，档次定义了语法（即算法），级别定义了各种参数（如分辨率、帧率和比特率等）。位于主级别的主要档次的目标应用是SDTV，而位于高级别的主要档次的目标应用为HDTV。

13.2.1　级别

MPEG-2支持4种级别，这4种级别分别为给定的档次定义了不同的分辨率、帧率、编码速率等。

1. 低级别（LL）

低级别（Low Level）中，MPEG-2约束参数比特流（CPB）支持的最大帧率为每秒30帧，最大分辨率为352×288，最大比特率为4Mbit/s。

2. 主级别（ML）

主级别（Main Level）中，MPEG-2约束参数比特流（CPB）支持的最大帧率为每秒30帧，最大分辨率为720×288，目标应用为SDTV。最大比特率为15~20Mbit/s。

3. 高1440级别

高1440级别（High 1440 Level）支持的最大帧率为每秒60帧，最大分辨率为1440×1080，其目标应用为HDTV。最大比特率为60~80Mbit/s。

4. 高级别（HL）

高级别（High Level）支持的最大帧率为每秒60帧，最大分辨率为1920×1088，其目标应用为HDTV。最大比特率为80~100Mbit/s。

13.2.2　档次

MPEG-2支持6个档次，这6个档次定义了各自所采用的编码语法（算法）。表13-1~表13-8说明了MPEG-2所允许的档次与级别的各种不同组合。

表13-1　MPEG-2可接受的级别与档次的组合

级别	档次						
	非可分级的				可分级的		
	简单	主要	多视角	4:2:2	SNR	空域	高
高	–	是	–	是	–	–	是
高1440	–	是	–	–	–	是	是
主	是	是	是	是	是	–	是
低	–	是	–	–	是	–	–

表13-2　部分MPEG-2档次约束

约束	档次						
	非可分级的				可分级的		
	简单	主要	多视角	4:2:2	SNR	空域	高
色度格式	4:2:0	4:2:0	4:2:0	4:2:0或者4:2:2	4:2:0	4:2:0	4:2:0或者4:2:2

（续）

约束	档次						
	非可分级的				可分级的		
	简单	主要	多视角	4：2：2	SNR	空域	高
图像类型	I，P	I，P，B	I，P，B	I，P，B	I，P，B	I，P，B	I，P，B
可分级性模式	–	–	时域	–	SNR	SNR或者空域	SNR或者空域
intra dc精度（位）	8，9，10	8，9，10	8，9，10	8，9，10，11	8，9，10	8，9，10	8，9，10，11
序列可分级性扩展	否	否	是	否	是	是	是
图像空域可分级性扩展	否	否	否	否	否	是	是
图像时域可分级性扩展	否	否	是	否	否	否	否
重复的第一个字段	受约束的		不受约束的	受约束的	不受约束的		

表13-3 MPEG-2可分级性档次允许的层次编号

级别	层次最大编号	档次			
		SNR	空域	高	多视角
高	所有层（基本层+增强层） 空域增强层 SNR增强层 时域备用层	–	–	3 1 1 0	2 0 0 1
高1440	所有层（基本层+增强层） 空域增强层 SNR增强层 时域备用层	–	3 1 1 0	3 1 1 0	2 0 0 1
主	所有层（基本层+增强层） 空域增强层 SNR增强层 时域备用层	2 0 1 0	–	3 1 1 0	2 0 0 1
低	所有层（基本层+增强层） 空域增强层 SNR增强层 时域备用层	2 0 1 0	–	–	2 0 0 1

表13-4 各个档次的若干MPEG-2视频解码器需求

档次	档次			基本解码器的各级别对应档次
	基本层	增强层1	增强层2	
SNR	4：2：0	SNR，4：2：0	–	位于同一级别的主要档次
	4：2：0	SNR，4：2：0	–	位于同一级别的主要档次
空域	4：2：0 4：2：0 4：2：0	空域，4：2：0 SNR，4：2：0 SNR，4：2：0	– 空域，4：2：0 SNR，4：2：0	位于下一级别的主要档次
高	4：2：0或4：2：2 4：2：0 4：2：0或4：2：2	– SNR，4：2：0 SNR，4：2：2	– – –	位于同一级别的主要档次

（续）

档次	档次			基本解码器的各级别对应档次
	基本层	增强层1	增强层2	
高	4：2：0	空域，4：2：0	–	位于下一级别的高档次
	4：2：0或4：2：2	空域，4：2：2	–	
	4：2：0	SNR，4：2：0	空域，4：2：0或4：2：2	
	4：2：0或4：2：2	SNR，4：2：2	空域，4：2：2	
	4：2：0	空域，4：2：0	SNR，4：2：0或4：2：2	
	4：2：0	空域，4：2：2	SNR，4：2：2	
	4：2：2	空域，4：2：2	SNR，4：2：2	
多视角	4：2：0	时域，4：2：0	–	位于同一级别的主要档次

表13-5 MPEG-2分辨率上限和时域参数。“增强层”参数适用于单层或者SNR可分级性编码情况

级别	空域分辨率层	参数	档次					
			简单	主要	多视角	4：2：2	SNR/空域	高
高	增强层	每行像素数 每帧行数 每秒帧数	–	1920 1088 60	1920 1088 60	1920 1088 60	–	1920 1088 60
	低层	每行像素数 每帧行数 每秒帧数	–	–	1920 1088 60	–	–	960 576 30
高1440	增强层	每行像素数 每帧行数 每秒帧数	–	1440 1088 60	1440 1088 60	–	1440 1088 60	1440 1088 60
	低层	每行像素数 每帧行数 每秒帧数		–	1440 1088 60	–	720 576 30	720 576 30
主	增强层	每行像素数 每帧行数 每秒帧数	720 576 30	720 576 30	720 576 30	720 608 30	720 576 30	720 576 30
	低层	每行像素数 每帧行数 每秒帧数	–	–	720 576 30	–	–	352 288 30
低	增强层	每行像素数 每帧行数 每秒帧数	–	352 288 30	352 288 30	–	352 288 30	–
	低层	每行像素数 每帧行数 每秒帧数	–	–	352 288 30	–	–	–

注：上述级别和档次中原来指定为每帧最大行数为1152改为每帧1088行。

表13-6 MPEG-2主要档次的级别和分辨率示例

级别	最大比特率(Mbit/s)	典型有效分辨率	帧率(Hz)[2] 23.976p	24p	25p	29.97p	30p	50p	59.94p	60p	25i	29.97i	30i
高	80(HP为100)(4:2:2档次为300)	1920×1080[1]	×	×	×	×	×				×	×	×
高1440	60(HP为80)	1280×720	×	×	×	×	×	×	×	×			
		960×1080[1]	×	×	×	×	×				×	×	×
		1280×1080[1]	×	×	×	×	×				×	×	×
		1440×1080[1]	×	×	×	×	×				×	×	×
主	15(HP为20)(4:2:2档次为50)	352×480	×	×		×	×		×	×		×	×
		352×576		×	×			×			×		
		480×480	×	×		×	×		×	×		×	×
		544×480	×	×		×	×		×	×		×	×
		544×576		×	×			×			×		
		640×480	×	×		×	×		×	×		×	×
		704×480,720×480	×	×		×	×		×	×		×	×
		704×576,720×576		×	×			×			×		
低	4	320×240	×	×		×	×		×	×		×	×
		352×240	×	×		×	×		×	×		×	×
		352×288		×	×			×			×		

注:1. 视频编码系统要求对于隔行图像有效扫描线数必须为32的倍数,对于逐行图像有效扫描线数必须是16的倍数。因此,对于1080线的隔行格式视频,视频编码器和解码器实际上必须采用1088行扫描线。多出的8线为没有实质内容的"虚拟"线,设计者之所以选择虚拟数据,是为了简化实现。多出的8线总是编码图像的最后8线。这些虚拟线并不携带有用信息,只是所需要传输的数据稍微有所增加。

2. p=逐行,i=隔行。

表13-7 MPEG-2 Y采样率上限(M采样点/秒)。"增强层"参数适用于单层或者SNR可分级编码的情况

级别	空间分辨率层	档次 简单	主要	多视角	SNR/空域	高	4:2:2
高	增强层	–	62.668800	62.668800	–	62.668800 (4:2:2) 83.558400 (4:2:0)	62.668800
	低层	–	–	62.668800	–	14.745600 (4:2:2) 19.660800 (4:2:0)	–
高1440	增强层	–	47.001600	47.001600	47.001600	47.001600 (4:2:2) 62.668800 (4:2:0)	–

（续）

级别	空间分辨率层	档次					
		简单	主要	多视角	SNR/空域	高	4：2：2
高1440	低层	–	–	47.001600	10.368000	11.059200 （4：2：2） 14.745600 （4：2：0）	–
主	增强层	10.368000	10.368000	10.368000	10.368000	11.059200 （4：2：2） 14.745600 （4：2：0）	11.059200
	低层	–	–	10.368000	–	3.041280 （4：2：0）	–
低	增强层	–	3.041280	3.041280	3.041280	–	–
	低层	–	–	3.041280	–	–	–

表13-8　MPEG-2比特率上限（Mbit/s）

级别	档次					
	非可分级的				可分级的	
	简单	主要	多视角	4：2：2	SNR/时域	高
高	–	80	130（所有层） 80（基本层）	300	–	100（所有层） 80（中间层＋基本层） 25（基本层）
高1440	–	60	100（所有层） 60（基本层）	–	60（所有层） 40（中间层＋基本层） 15（基本层）	80（所有层） 60（中间层＋基本层） 20（基本层）
主	15	15	25（所有层） 5（基本层）	50	15（所有层） 10（基本层）	20（所有层） 15（中间层＋基本层） 4（基本层）
低	–	4	8（所有层） 4（基本层）	–	4（所有层） 3（基本层）	–

1. 简单档次

无B帧的主要档次即为简单档次（Simple Profile，SP），目标应用为软件应用和可能的有线电视。

2. 主要档次

大多数MPEG-2解码器芯片都支持主要档次（Main Profile，MP），主要档次应该满足90%的消费类SDTV和HDTV应用。表13-6列出了主要档次支持的典型分辨率。

3. 多视角档次

通过采用已有的MPEG-2工具，多视角档次（Multiview Profile，MVP）可能能够解码同一场景、角度有稍微差别的两个摄像机拍摄的视频。

4. 4：2：2档次

4：2：2档次（4：2：2 Profile，422P）即为人们熟知的“演播室档次（studio profile）”，422P采用4：2：2的YCbCr替代4：2：0的YCbCr，此外，相对主要档次，422P增加了最大比特率到

50Mbit/s（对于高档次，这个值为300Mbit/s），从而支持专业级视频SDTV和HDTV需求。

578

5. SNR和空域档次

SNR和空域档次（SNR and Spatial Profile）增加了对SNR可分级性和空域可分级性的支持。

6. 高档次

高档次（High Profile，HP）目标应用为专业级视频HDTV应用。

13.2.3 可分级性

MPEG-2 SNR、空域和高档次支持四种可分级的操作模式。这些模式将MPEG-2视频分成不同的层次，从而能够区分视频数据的优先次序。采用分级性将导致效率下降近2dB（或者需要多出约30%比特），因此可分级性并不被普遍采用。

1. SNR可分级性

SNR可分级模式的目标应用是那些需要提供多个质量层次的应用。在该模式中，所有层都有相同的空间分辨率。基本层提供基本的视频质量，增强层通过为基本层DCT系数提供精细数据而提升视频质量。

2. 空域可分级性

空域可分级性对于同播（simulcasting）是非常有用的，在该模式中每一个层都有一个不同的空间分辨率。基本层提供基本的空间分辨率和时域速率。增强层采用对基本层进行空间插值来增加空域分辨率。例如，基本层可能实现352×240分辨率视频，利用增强层可生成704×480分辨率的视频。

3. 时域可分级性

时域可分级模式使得可以从低时域速率迁移到更高的时域速率系统。基本层提供基本的时域速率。增强层采用相对于基本层的时域预测。基本层和增强层可以组合在一起从而产生一个全时域速率输出。所有层都有相同的空域分辨率和色度格式。在增强层出现差错的情况下，基本层可以用来进行差错掩蔽（concealment）。

4. 数据分区

数据分区模式的目标是在ATM网络中进行信元丢失恢复（resilience）。在该模式中，将64个量化的变换系数分成两个比特流。高优先级比特流包含关键的低频DCT系数和伴随信息，如头和运动矢量等。低优先级比特流携带可以增加细节的高频DCT系数。

13.2.4 传输和节目流

MPEG-2系统标准定义了两种方法来复用音频、视频和其他数据到一个适合于传输和存储的格式。

节目流是为很少出现差错的应用而设计的。它包含了音频、视频和数据比特流（也称作基本比特流），并将这些比特流合并到一个单独的比特流中。节目流，也就是每一个基本比特流，可能是固定或者可变比特率。DVD和SVCD采用节目流，在与视频和音频流交织的私有数据流中携带DVD和SVCD相关的数据。

579
584

传输流是为较容易出现数据丢失的应用而设计的，它采用长度为188字节的固定大小包。传输流同样包含了音频、视频和数据比特流，并且将这些比特流都合并到一个单独的比特流，传输流可以携带多个节目。ARIB、ATSC、DVB和OpenCable™标准都采用传输流。

传输流和节目流都是基于通用的包结构，便于通用的解码器实现和转换。这两种流都是设计支持大量已知和未来可以预见的应用，且保持灵活性。

13.3　视频编码层

13.3.1　YCbCr颜色空间

MPEG-2采用YCbCr颜色空间，支持4：2：0、4：2：2和4：4：4采样。4：2：2和4：4：4采样选项在4：2：0采样上提高了色度分辨率，从而获得更好的图像质量。

图3-8~图3-10给出了MPEG-2的4：2：0采样结构，图3-2和图3-3分别给出了MPEG-2的4：2：2和4：4：4采样结构。

13.3.2　编码图像类型

MPEG-2中有三种类型的编码图像。I（intra）图像是作为独立静态图像编码的场或帧。I图像在视频流中提供可以随机存取的随机存取点。通常，在1秒钟的视频中约出现2次I图像。此外，在场景发生突变的地方也应该采用I图像。

P（predicted）图像是由与其最近邻的前一个I图像或P图像预测而得到的场或帧（即所谓的前向预测处理），如图13-1所示。通过采用运动补偿技术，P图像提供比I图像更好的压缩效率，而且也能够作为B图像和接下来的P图像的参考图像。

B（bi-directional）图像是采用最邻近的前后两个I图像或P图像作为参考，进行双向预测得到的场或帧，如图13-1所示。B图像提供最高的压缩效率，并且通过利用前后两个图像的均值减少了噪声。通常，I图像或P图像之间有2个B图像。

MPEG-2中不支持D（DC）图像，除了在解码过程中支持与MPEG-1的向后兼容性之外。

GOP是用于辅助随机存取和编辑的一个或多个编码图像序列。在编码过程中GOP值是可以配置的，GOP值越小，对运动的响应效果越好（因为与I图像靠得越近），但压缩效率却下降了。

在编码流中，一个图像组必须以I图像作为开始，随后可以按任意顺序连接任意个I图像、P图像或B图像。在显示顺序中，图像组必须以I图像或B图像开始，并以I图像或P图像结束。因此，最小的图像组是单个I图像，而最大图像组的大小是无限的。

每个图像组应该独立于其他任何图像组进行编码。然而，这种情况只有在第一个I图像前没有B图像才有可能，否则，只能进行逆向运动补偿。于是形成了开放和封闭两种图像组格式。封闭图像组不需要之前的图像组做运动补偿就能完成解码。而采用broken_link标志标识的开放图像组[意味着在GOP头之后的第一个B图像（如果有）是紧跟在第一个I图像之后]可能无法正确地解码（因此也就不会显示），因为用于预测的参考图像由于编辑原因而不能获得。

585

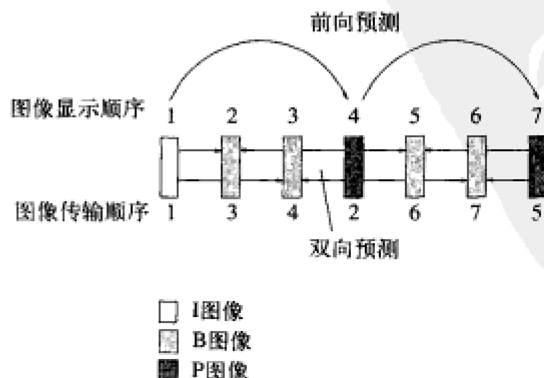


图13-1　MPEG-2中的I、P和B图像。某些图像可能不按顺序传输，从而使插值过程变得复杂，因此需要MPEG解码器对图像进行重新排序。箭头表示帧之间的依赖关系

13.3.3 运动补偿

由于引入了场（field），MPEG-2的运动补偿变得更为复杂。采用运动补偿对一个宏块进行解压之后，它既包含了参考宏块和解码宏块之间的空域差值（运动矢量），也包含了内容差值（误差项）。

两个主要的预测类别分别是场和帧。在场图像中，仅采用场预测。在帧图像中，既可以采用场预测，也可以采用帧预测（在宏块层决定具体采用何种预测）。

MPEG-2的运动矢量总是采用半像素精度进行编码，而MPEG-1既支持半像素也支持全像素精度。

1. 16×8运动补偿选项

在该选项中，采用每个宏块两个运动矢量（对于B图像则为4个），其中一个用于宏块上部的16×8区域，另一个用于宏块下部的16×8区域。该模式仅用于场图像。

2. 对偶运动补偿选项

该模式仅用于在预测和参考场或帧之间没有B图像的P图像。采用一个运动矢量，同时还有一个小的差分运动矢量。所有必要的预测都是由这个矢量推导出来的。 586

13.3.4 宏块

MPEG-2中有三种类型的宏块。

4：2：0宏块（图13-2）由4个Y块、1个Cb块和1个Cr块组成。图中给出了块的排列次序。

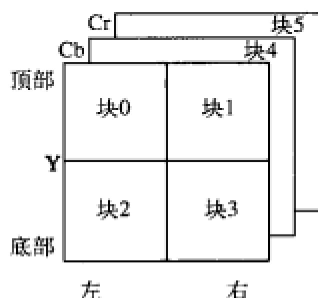


图13-2 MPEG-2 4：2：0宏块结构

4：2：2宏块（图13-3）由4个Y块、2个Cb块和2个Cr块组成。图中给出了块排列次序。

4：4：4宏块（图13-4）由4个Y块、4个Cb块和4个Cr块组成。图中给出了块的排列次序。

P图像中的宏块采用最近的前一个I或P图像作为参考编码而得到，因此导致下面两种可能的编码模式：

- ❑ 内部编码（intra-coding），没有运动补偿；
- ❑ 前向预测（forward prediction），最近的前一个I或P图像作为参考。

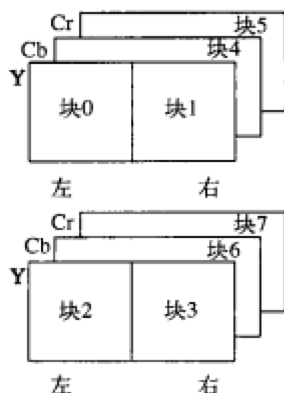


图13-3 MPEG-2 4：2：2宏块结构

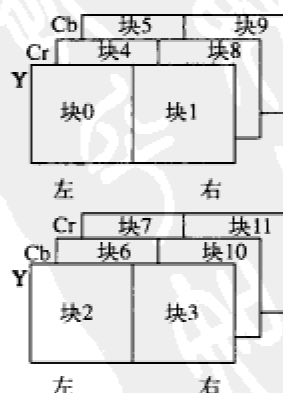


图13-4 MPEG-2 4：4：4宏块结构

B图像中的宏块采用最近的前一个和后一个I或P图像作为参考而编码得到，因此导致下面四

种可能的编码模式：

- 内部编码（intra-coding），没有运动补偿；
- 前向预测（forward prediction），最近的前一个I或P图像作为参考；
- 后向预测（backward prediction），最近的后一个I或P图像作为参考；
- 双向预测（bi-directional prediction），采用两个图像作为参考——最近的前一个I或P图像和最近的后一个I或P图像。

13.3.5 I图像

1. 宏块

I图像中有10种类型的宏块，如表13-27所示。

如果表13-27中的[macroblock quant]列值为1，量化器刻度将会传送。对于其他宏块类型，DCT修正将采用前一个值来编码，编码结果作为该宏块的量化器刻度。

如果表13-27中的[coded pattern]列值为1，长度为6位的编码块模式将被作为可变长编码传送。编码块模式将指示解码器在4：2：0宏块的6个块中哪一个块已经编码（标示为1），哪一个没有编码（标示为0）。表13-32列出了分配给63种可能组合的码字。当所有6个块都没有编码时，将没有码字，这由宏块类型来指示。对于4：2：2和4：4：4宏块，分别有附加的2位或者6位用来扩展编码块模式。

2. DCT

每一个8×8块（输入采样点或者预测误差项）由一个8×8 DCT处理，得到一个8×8的垂直和水平频率系数的块。

输入采样点值为0~255，DCT变换之后，获得的DC系数范围为0~2040，AC系数为−2040~2047。

由于空域和SNR可分级性，非内部块（在非内部宏块内的块）的出现同样是可能的。非内部块系数表示的是采样值之间的差值，而不是实际采样值。它们是通过从当前宏块的运动补偿值中抽取出前一个图像中的运动补偿值而获得的。

587

3. 量化

频率系数8×8块被均匀地量化，从而限制了量化值的数目。量化器步长刻度由量化矩阵和量化器刻度推导而得到，并且对于不同的系数可能有不同的值，在宏块之间也可能发生变化。

由于人的眼睛对于大亮度区域较为敏感，因此，DC系数的量化器步长刻度可选择8、9、10或11位的精度。量化的DC系数由DC系数除以8、4、4或1之后取整到最近的整数而确定。

AC系数采用两个量化矩阵量化：其中一个用于内部宏块，另一个用于非内部宏块。当采用4：2：2或者4：4：4数据时，对于Y和CbCr数据可能采用不同的矩阵。每一个量化矩阵都有一个默认的值集，该默认的值集可以被覆盖。

如果表13-27中的[macroblock quant]列值为1，量化器刻度将被传送。对于其他宏块类型，DCT修正将采用前一个值来编码，编码结果作为该宏块的量化器刻度。

4. Z形扫描

如图7-57和图7-58所示，Z形扫描从DC分量开始，按照频率递增的顺序将量化频率系数排列成线性流。这将产生长游程的非零系数。

5. 量化DC系数编码

DC系数量化后，将采用无失真编码方式对其进行编码。

宏块中的Y块编码按照图13-2所示的顺序进行。块4的DC值将作为下一个宏块中块1的DC值的预测值。在每个条带的开头，当一个宏块被跳过或者当一个非内部宏块被编码时，DC预测值

被置为128（DC精度为8位）、256（DC精度为9位）、512（DC精度为10位）或者1024（DC精度为11位）。

每个Cb、Cr块的值采用前一个宏块的对应块的DC值作为预测值进行编码，在每个条带开头，当一个宏块被跳过或者当一个非内部宏块被编码时，DC预测值被置为128（DC精度为8位）、256（DC精度为9位）、512（DC精度为10位）或者1024（DC精度为11位）。

然而，一个普遍的实现是将DC预测值设置为0，并且将内部块DC项中心值设置为0附近，而取代50%灰度值。因此，解码器仅需要处理量化器（有一个乘法器可以用来重建正确值），而不是剖析器（通常不接触数据，且没有乘法器）内的不同内部DC精度。

6. 量化AC系数编码

AC系数量化后，将按照图7-57或图7-58所示的顺序扫描，采用游程（run-length）和非零值（level）编码。如图7-57和图7-58所示，扫描从位置1开始，因为在位置0的DC系数单独编码。

游程和非零值采用如表13-36和表13-37所示方式编码。s位表示该非零值的符号，0为正，1为负。对于内部块，既可以采用表13-36，也可以采用表13-37，具体采用哪个表将由比特流中的intra_vlc_format来决定。对于非内部块，只能采用表13-36。

588 ≀ 589

对于表13-36和表13-37中没有列出的游程－非零值组合的情况，引入了一个转义序列，由一个转义码（ESC）后面接着如表13-38和表13-39所示的游程和非零值编码。

当最后一个DCT系数编码完成后，添上一个EOB符号，以指示解码器在这个8×8块中已经没有其他量化系数了。

13.3.6 P图像

1. 宏块

如表13-28所示，由于运动补偿附加的复杂性，P图像内有一共26种类型的宏块。

如果比特流中的macroblock_address_increment参数大于1，不编码宏块（skipped macroblock）将会出现。对于P场图像，解码器将从与被预测场同奇偶性的场进行预测，运动矢量预测值被设置为0。对于P帧图像，解码器设置运动矢量预测值和运动矢量都为0。

如果表13-28中的[macroblock quant]项为1，需要传送量化器刻度。对于其他宏块类型，采用先前的量化器刻度对DCT修正进行编码。

如果表13-28中的[motion forward]列值为1，水平和垂直前向运动矢量应当相继传输。

如果表13-28中的[coded pattern]列值为1，6位编码块模式（coded block pattern）将作为变长编码传输。它会指示解码器当前宏块的6个块中哪些已经编码（设为1），哪些没有编码（设为0）。表13-32列出63种可能组合对应的码字。如果没有块被编码，就没有码字。编码块模式由宏块类型决定。对于P帧、B帧中的帧内编码宏块，编码块模式将不需要传输，且假定为63（表示所有块都已被编码）。对于4：2：2和4：4：4宏块，分别有附加的2位或者6位用来扩展编码块模式。

2. DCT

内部块AC系数采用与I图像中的内部块AC系数相同的方式进行变换。内部块DC系数变换方式则不同，预测值设置为1024，除非前一个块是采用帧内编码方式进行编码的。

非内部块系数表示的是采样值之间的差值，而不是实际采样值，它们是通过从当前宏块的运动补偿值中抽取出前一个图像中的运动补偿值而获得的。DC值是没有预测的。

输入采样点值范围为－255~＋255，DCT变换之后，获得的AC系数范围为－2000~＋2000。

3. 量化

P图像中的内部块采用与I图像中内部块相同的方式进行量化。

非内部块采用量化器刻度和非内部量化矩阵进行量化。AC和DC系数采用相同的方式进行量化。

4. 内部块编码

P图像中的内部块采用与I图像中内部块相同的方式进行编码。在处理DC系数时稍微有点不同，即在P图像中预测值设为128，除非前一个块是内部编码块。

5. 非内部块编码

590 编码块模式（CBP）用来指定哪一个块有系数数据。除了在非内部块中DC系数与AC系数采用系统的编码方式之外，非内部块编码方式与内部块编码方式类似。

13.3.7 B图像

1. 宏块

根据后向运动补偿附加的复杂程度不同，B图像中有34种宏块类型，如表13-29所示。

对于B场图像，解码器将从与被预测场同奇偶性的场进行预测。预测方向（前向、后向或者双向）与前一个宏块的预测方向相同，运动矢量预测值不受影响，且运动矢量将从适当的运动矢量预测值中得到。对于B帧图像，预测方向（前向、后向或者双向）与前一个宏块的预测方向相同，运动矢量预测值不受影响，且运动矢量将从适当的运动矢量预测值中得到。

如果表13-29中的[macroblock quant]列值为1，将传输量化器刻度。而对于剩下的宏块类型，则采用前一个量化器刻度对DCT修正进行编码。

如果表13-29中的[motion forward]列值为1，将要相继传输水平和垂直前向运动矢量。如果表13-29中的[motion backward]列值为1，将要相继传输水平和垂直后向运动矢量。如果前向后向运动类型都出现了，运动矢量按照以下顺序传输：水平前向→垂直前向→水平后向→垂直后向。

如果表13-29中的[coded pattern] 列值为1，6位长的编码块模式将采用变长编码方式传输，并指示解码器当前宏块的6个块中哪一块已经被编码（设为1），哪一块还没有被编码（设置为0）。表13-32列出了63种可能组合对应的码字。如果还没有块被编码，就没有码字。编码块模式由宏块类型决定。对于P帧、B帧中的帧内编码宏块，编码块模式不需要传输，且假定为63（表示所有块都已经被编码）。对于4：2：2和4：4：4宏块，分别有附加的2位或者6位用来扩展编码块模式。

2. 编码

块的DCT系数被转换成量化的系数，并且采用与P图像中相同的方式进行编码。

13.4 视频比特流

图13-5对视频比特流进行了举例说明，从上到下总共7个层次，各层分别是：视频序列、序列头、图像组（GOP）、图像、条带、宏块（MB）、块。

以下扩展可以用来支持不同级别的能力：序列扩展、序列显示扩展、序列可分级扩展、图像

591 编码扩展、量化矩阵扩展、图像显示扩展、图像时域可分级扩展、图像空域可分级扩展。

如果视频序列的第一个序列头不是跟在扩展起始码（0x00001B5）之后，视频比特流必须符合MPEG-1视频比特流规范。

对于MPEG-2视频比特流，扩展起始码（0x000001B5）和序列扩展之后必须跟每一个序列头。

注意，起始码（0x000001xx）必须是字节对齐的，其实现方式是在起始码之前插入0~7个0位。

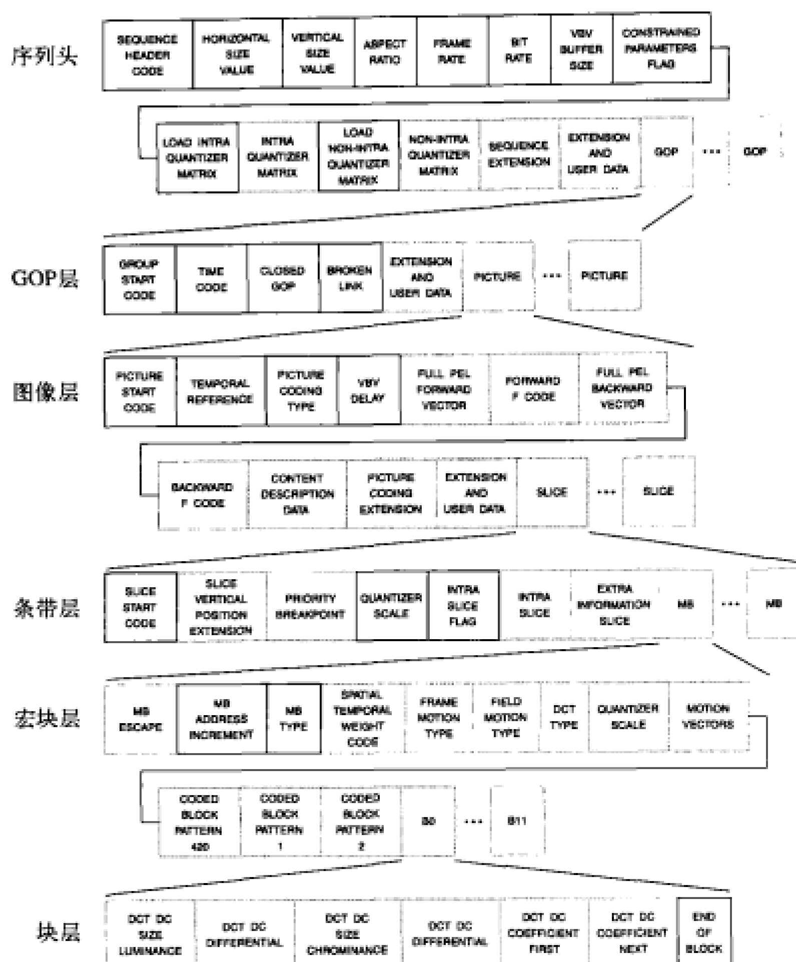


图13-5 MPEG-2视频比特流层次结构。没有给出标记和保留位

13.4.1 视频序列

sequence_end_code

长度为32位的字段，值为0x000001B7，标志一个视频序列的结束。

13.4.2 序列头

一般来说，平均每秒钟应出现两次序列头。序列头结构如图13-5所示。如果其后没有跟一个序列扩展，那么比特流符合MPEG-1规范。

1. sequence_header_code

长度为32位的字符串，值为0x000001B3，表示序列头的开始。

2. horizontal_size_value

这是一个12位最低有效位，表示Y分量可见部分的宽度（像素）。14位值的两个最高有效位在horizontal_size_extension中指定了。零值是不允许的。

3. vertical_size_value

这是一个12位最低有效位，表示Y分量可见部分的高度（扫描线）。14位值的两个最高有效位

在vertical_size_extension中指定了。零值是不允许的。

4. aspect_ratio_information

这是一个4位的码字，指示如表13-9所示的像素宽高比（SAR）或者显示宽高比（DAR）。

如果sequence_display_extension没有出现，SAR由下面的公式确定：

SAR＝DAR × (horizontal_size/vertical_size)

如果sequence_display_extension出现了，SAR由下面的公式确定：

SAR＝DAR × (display_horizontal_size/display_vertical_size)

5. frame_rate_code

这是一个4位的码字，指示如表13-10所示的帧率。

实际帧率由下面的公式确定：

frame_rate＝frame_rate_value × (frame_rate_extension_n＋1)/(frame_rate_extension_d＋1)

592～593

当指定表13-10中的某一个条目时，frame_rate_extension_n和frame_rate_extension_d均为00。如果progressive_sequence为1，两帧到达解码器输出点之间的时间为帧率（frame_rate）的倒数。如果progressive_sequence为0，两帧到达解码器输出点之间的时间为帧率（frame_rate）的倒数的一半。

表13-9 MPEG-2 aspect_ratio_information码字

SAR	DAR	编　码
禁止	禁止	0000
1.0000	–	0001
–	3/4	0010
–	9/16	0011
–	1/2.21	0100
–	保留	0101
–		0110
–		0111
–		1000
–		1001
–		1010
–		1011
–		1100
–		1101
–		1110
–		1111

表13-10 MPEG-2 frame_rate_code码字

帧/秒	编　码
禁止	0000
24/1.001	0001
24	0010
25	0011
30/1.001	0100
30	0101
50	0110
60/1.001	0111
60	1000
保留	1001
保留	1010
保留	1011
保留	1100
保留	1101
保留	1110
保留	1111

6. bit_rate_value

30位二进制数中的18位最低有效位，12位最高有效位在bit_rate_extension中。bit_rate_value表示比特流的比特率，以400bit/s为单位度量，不足时取上界。该值不允许为0。对于ATSC标准，值必须小于等于48500_D（对于高数据率模式为小于等于97000_D）。对于OpenCable™标准中的64QAM系统，值必须小于等于67500_D（对于256QAM系统为小于等于97000_D）。

7. marker_bit

标记位，值恒为1。

8. vbv_buffer_size_value

18位二进制数中的10位最低有效位。8位最高有效位在vbu_buffer_extension中。vbv_buffer_size

定义了解码器对视频序列解码时，视频缓存校验器（Video Buffering Verifier）需要的最小存储容量。按照如下定义：

$B = 16 \times 1024 \times \text{vbv_buffer_size}$

对于ATSC和OpenCable™标准，该值必须小于等于488_D。

9. constrained_parameters_flag

此位设为0，因为它对于MPEG-2没有任何意义。

10. load_intra_quantizer_matrix

如果其后紧接intra_quantizer_matrix，此位设为1。如果设为0，如下默认值将用于内部块（Y和CbCr），直到下一个序列头或者quant_matrix_extension出现：

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

11. intra_quantizer_matrix

可选的64个8位值列表，用来替换当前的帧内量化器值。不允许为0。intra_quant[0,0]通常为8。在下一个序列头或者quant_matrix_extension出现之前，这些数值一直有效。对于4：2：2和4：4：4数据格式，新的值将用于Y和CbCr内部矩阵，除非装载了一个不同的CbCr内部矩阵。

12. load_non_intra_quantizer_matrix

如果其后紧跟non_intra_quantizer_matrix，此位设为1。如果设为0，在下一个序列头或者quant_matrix_extension出现之前，如下默认值将用于内部块（Y和CbCr）：

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

13. non_intra_quantizer_matrix

可选的64个8位值列表，用来替换当前的非帧内量化器值。不允许为0。在下一个序列头或者quant_matrix_extension出现之前，这些数值一直有效。对于4：2：2和4：4：4数据格式，新的值将用于Y和CbCr非内部矩阵，除非装载了一个不同的CbCr非内部矩阵。

594 ~ 595

13.4.3 用户数据

1. user_data_start_code

可选的32位字符串，值为0x000001B2，表示user_data（用户数据）的开始。在探测到下一个起始码之前，user_data一直持续。

2. user_data

长度为有$n \times 8$位，当user_data_start_code出现时才出现。user_data不能包含超过连续23个0的字符串。

13.4.4 序列扩展

序列扩展只可能在序列头之后出现，其结构如图13-6所示。

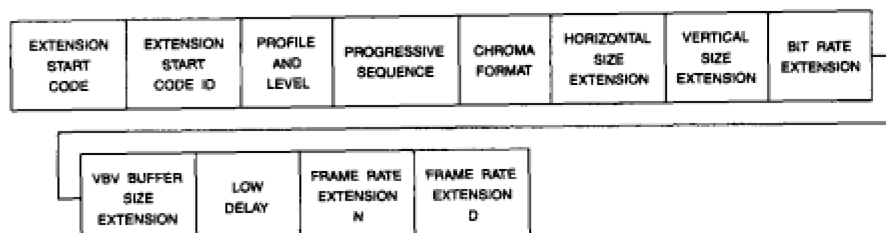


图13-6 MPEG-2序列扩展结构。没有给出标记位

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示除MPEG-1之外的扩展数据的开始。

2. extension_start_code_ID

长度为4位的字段，值为0001，表示序列扩展的开始。对于MPEG-2视频比特流，每一个序列头都必须跟一个序列扩展。

3. profile_and_level_indication

长度为8位的字段，用于定义档次和级别，如表13-11所示。

位7表示转义位，位6~4表示档次ID，

位3~0表示级别ID。

表13-11 MPEG-2 profile_and_level_indication码字

档　次	档次ID编码	级　别	级别ID编码
保留	000	保留	0000
高	001	保留	0001
空域可分级	010	保留	0010
SNR可分级	011	保留	0011
主要	100	高	0100
简单	101	保留	0101
保留	110	高1440	0110
保留	111	保留	0111
		主要	1000
		保留	1001
		低	1010
		保留	1011
		保留	1100
		保留	1101
		保留	1110
		保留	1111

4. progressive_sequence

1位标志位，值为1时指示仅逐行图像出现，值为0时表示帧和场图像都可能出现，且帧图像可能是逐行的，也可能是隔行的。对于SVCD标准，该值必须为0。

5. chroma_format

2位的码字，表示CbCr格式，如表13-12所示。对于ASTC和OpenCable™标准，该值必须为01。

6. horizontal_size_extension

horizontal_size的两位最高有效位。对于ASTC和OpenCable™标准，该值必须为00。

7. vertical_size_extension

vertical_size的两位最高有效位。对于ASTC和OpenCable™标准，该值必须为00。

8. bit_rate_extension

bit_rate的12位最高有效位。对于ASTC和OpenCable™标准，该值必须为0000 0000 0000。

9. marker_bit

标记位，值恒为1。

10. vbv_buffer_size_extension

vbv_buffer_size的8位最高有效位。对于ASTC和OpenCable™标准，该值必须为0000 0000。

11. low_delay

1位标志位，该位为1时表示没有B图像出现，因此没有帧重排序时延。对于SVCD，该值必须为0。 596

12. frame_rate_extension_n

参看frame_rate_code关于这2位的二进制值。对于ASTC和OpenCable™标准，该值必须为00。

13. frame_rate_extension_d

参看frame_rate_code关于这5位的二进制值。对于ASTC和OpenCable™标准，该值必须为00000。

13.4.5 序列显示扩展

这个可选的扩展仅可能在序列扩展之后出现，其结构如图13-7所示。

EXTENSION START CODE	EXTENSION START CODE ID	VIDEO FORMAT	COLOR DESCRIPTION	COLOR PRIMARIES	TRANSFER CHARACTERISTICS	MATRIX COEFFICIENTS	DISPLAY HORIZONTAL SIZE	DISPLAY VERTICAL SIZE

图13-7 MPEG-2序列显示扩展结构。没有显示出标记位

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为0010，表示序列显示扩展的开始。由该扩展提供的信息并不影响解码过程，且可以被忽略。该扩展使得能够尽可能正确地显示解码的图像。

3. video_format

长度为3位的码字，指示编码之前的图像源格式，如表13-13所示。对于ASTC和OpenCable™标准，该值必须为000。

表13-12 MPEG-2 chroma_format码字

色度格式	编　码
保留	00
4：2：0	01
4：2：2	10
4：4：4	11

表13-13 MPEG-2 video_format码字

视频格式	编　码
分量	000
PAL	001
NTSC	010
SECAM	011
MAC	100
未指定	101
保留	110
保留	111

4. color_description

1位的标志位，值为1时表示在比特流中将出现color_primaries、transfer_characteristics和matrix_cofficients。

5. color_primaries

可选的8位码字，描述了源基色的色度坐标，如表13-14所示。如果sequence_display_extension没有出现，或者color_description＝0，表示必须采用默认值。

597～598

该信息可以在MPEG-2解码之后用来调整颜色处理过程，从而补偿显示的基色。

6. transfer_characteristics

可选的8位码字，描述了源图像的光电传输特性，如表13-15所示。如果sequence_display_extension没有出现，或者color_description＝0，表示必须采用默认值。

该信息可以在MPEG-2解码之后用来调整颜色处理过程，从而补偿显示的光电转换特性（gamma）。

表13-14 MPEG-2 color_primaries码字

Color Primaries	编　码	默认应用
禁止	0000 0000	
BT.709、SMTPE 274M、BT.1361、IEC 61966-2-4	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
BT.470系统M	0000 0100	30Hz SDTV
BT.470系统B、G	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
保留	0000 1000	
⋮	⋮	⋮
保留	1111 1111	

表13-15 MPEG-2_transfer_characteristic码字

光电传输特性	编　码	默认应用
禁止	0000 0000	
BT.709、SMPTE 274M、BT.1361	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
BT.470系统M	0000 0100	30Hz SDTV
BT.470系统B、G	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
线性	0000 1000	
对数（100：1范围）	0000 1001	
对数（316：1范围）	0000 1010	
IEC 61966-2-4	0000 1011	
BT.1361	0000 1100	
保留	0000 1101	
⋮	⋮	⋮
保留	1111 1111	

7. matrix_cofficients

可选的8位码字，描述了用于从R'G'B'推导YCbCr的系数，如表13-16所示。如果

sequence_display_extension没有出现，或者color_description＝0，表示必须采用默认值。

根据情况需要，该信息可以在MPEG-2解码之后用来选择合适的YCbCr到RGB转换的矩阵。

表13-16 MPEG-2 matrix_cofficients码字

矩阵系数	编码	默认应用
禁止	0000 0000	
BT.709、SMPTE 274M、 BT.1361、 IEC 61966-2-4（xvYCC709）	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
FCC	0000 0100	30Hz SDTV
BT.470系统B、G、I， IEC 61966-2-4（xvYCC601）	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
YCgCo	0000 1000	
保留	0000 1001	
⋮	⋮	⋮
保留	1111 1111	

8. display_horizontal_size

参看display_vertical_size，查看关于这个14位的二进制数的相关情况。

9. marker_bit

标记位，值恒为1。

10. display_vertical_size

14位的二进制数，与display_horizontal_size一起定义了显示的有效区域。如果显示区域小于解码的图像大小，只能显示图像的一部分。如果显示区域大于解码图像大小，图像将显示在显示区域的一部分上。

13.4.6 序列可分级扩展

这个可选的扩展仅可能在序列扩展之后出现，其结构如图13-8所示。

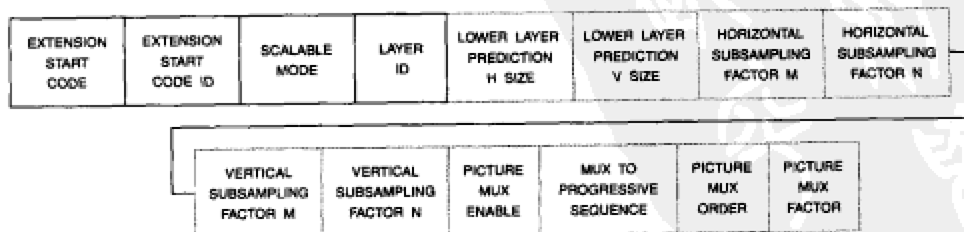


图13-8 MPEG-2序列可分级扩展结构。标记位没有给出

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为0101，表示序列可分级扩展的开始。该扩展定义了位视频比特流实现的可分级模式。如果sequence_scalable_extension没有出现在比特流中，就没有可分级性被采用。

599～601 除了在数据分区可分级模式中外，其他可分级层次结构的基本层都没有sequence_scalable_extension。

3. scalable_mode

两位的码字，表示视频序列的可分级类型，如表13-17所示。

表13-17 MPEG-2 scalable_mode码字

可分级模式	编　码
数据分区	00
空域可分级性	01
SNR可分级性	10
时域可分级性	11

4. layer_ID

4位的二进制数，标识在可分级层次结构中的层次。基本层ID为0000。在数据分区中，0000的layer_ID被分配给了分区层0，而0001的layer_ID被分配给了分区层1。

5. lower_layer_prediction_horizontal_size

可选的14位二进制数，仅当scalable_mode＝01时才出现。它表示用于预测的低层帧的水平大小。它包含了在低层比特流中的horizontal_size值。

6. marker_bit

标记位，值恒为1。它只有在scalable_mode＝01时才出现。

7. lower_layer_prediction_vertical_size

可选的14位二进制数，仅当scalable_mode＝01时才出现。它表示用于预测的低层帧的垂直大小。它包含了在低层比特流中的vertical_size值。

8. horizontal_subsampling_factor_m

可选的5位二进制数，仅当scalable_mode＝01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

9. horizontal_subsampling_factor_n

可选的5位二进制数，仅当scalable_mode＝01时才出现，该参数对于空域上采样过程有影响。602 取值不能为00000。

10. vertical_subsampling_factor_m

可选的5位二进制数，仅当scalable_mode＝01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

11. vertical_subsampling_factor_n

可选的5位二进制数，仅当scalable_mode＝01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

12. picture_mux_enable

可选的1位字段，仅当scalable_mode＝11时才出现。如果设置为1，picture_mux_order和picture_mux_factor参数将用于在显示之前的重复用（remultiplexing）。

13. mux_to_progressive_sequence

可选的1位标志位，仅当scalable_mode＝11且picture_mux_enable＝1时才出现。设置为1表示解码图像将在时域上复用，以生成用于显示的逐行序列。当时域复用将生成隔行序列时，该标志位需要设置为0。

14. picture_mux_order

可选的3位二进制数，仅当scalable_mode＝11时才出现。它定义了在第一个基本层图像之前的增强层图像数目。该参数将用于帮助解码器在显示之前以适当的方式重复用图像。

15. picture_mux_factor

可选的3位二进制数，仅当scalable_mode＝11时才出现。它表示在两个连续的基本层图像之

间的增强层图像数目。该参数将用于帮助解码器在显示之前以适当的方式重复用图像。

13.4.7 图像组（GOP）层

GOP头应该每两秒钟出现一次。每一个图像组的数据都由一个GOP头和跟随其后的图像数据组成，其结构如图13-5所示。DVD标准在这一层为隐藏字幕数据采用了用户数据扩展。

1. group_start_code

长度为32位的字符串，值为0x000001B8，表示一个图像组的开始。

2. time_code

表示时间码信息的25位字段，如表13-18所示。仅当帧率为30/1.001（29.97）Hz时，drop_frame_flag才可以被设置为1。

表13-18 MPEG-2 time_code字段

Timecode	取值范围	位 数
drop_frame_flag		1
time_code_hours	0~23	5
time_code_minutes	0~59	6
marker_bit	1	1
time_code_seconds	0~59	6
time_code_pictures	0~59	6

3. closed_gop

1位标志位，当图像组没有参考前一个图像组运动矢量进行编码时，该标志位设置为1。该位使得能够支持对压缩比特流的编辑。

4. broken_link

1位标志位，在解码过程中设置为0。在编辑过程中，当跟在图像组的第一个I帧之后的B帧不能正确地解码时，该标志位设置为1。 603

13.4.8 图像层

每一个图像的数据由一个图像头与跟随在之后的条带数据组成，其结构如图13-5所示。如果序列扩展出现了，每一个图像头将跟在图像编码扩展之后。

在这一层，通过用户数据扩展，有些实现使得帧与帧之间可以进行精准的高宽比信息切换。ATSC标准同样在这一层为CEA-708隐藏字幕数据采用了用户数据扩展。

1. picture_start_code

长度为32位的字符串，值为0x00000100。

2. temporal_reference

10位的二进制数，对于GOP中的第一帧，temporal_reference值为零。对于随后的显示顺序中每一帧，temporal_reference以1024为模依次递增（增量为1）。当某一帧以两场编码时，两个场的时域参考是一致的。

3. picture_coding_type

3位的码字，表示图像类型（I图像、P图像或者B图像），如表13-19所示。

表13-19 MPEG-2 picture_coding_type码字

图像类型	编 码
禁止	000
I图像	001
P图像	010
B图像	011
禁止	100
保留	101
保留	110
保留	111

4. vbv_delay

16位的二进制数，对于恒定比特率，在开始解码一个图像时，该参数将设置解码缓存的初始值，从而不会导致解码时发生上溢或者下溢的情况。对于ATSC和OpenCable™标准，除非vbv_delay值为0xFFFF，否则其值必须小于等于45000_D。 604

5. full_pel_forward_vector

可选的1位字段，没有用于MPEG-2，因此其值为0。仅当picture_coding_type＝010或者011时才出现。

6. forward_f_code

可选的3位字段，没有用于MPEG-2，因此其值为111。仅当picture_coding_type＝010或者011时才出现。

7. full_pel_backward_vector

可选的1位字段，没有用于MPEG-2，因此其值为0。仅当picture_coding_type＝011时才出现。

8. forward_f_code

可选的3位字段，没有用于MPEG-2，因此其值为111。仅当picture_coding_type＝011时才出现。

9. extra_bit_picture

1位标志位，当设置为1时，表示其后跟随content_description_data。当设置为0时，表示其后没有content_description_data。

10. content_description_data

如果extra_bit_picture＝1，该可选的可变长字段将会出现，其中每隔8位有一位值为1。

11. extra_bit_picture

1位标志位，值为0，且仅当content_description_data出现时才出现。

13.4.9　内容描述数据

可选的数据，且仅当图像头中的extra_bit_picture指示其出现时才出现。

1. data_type_upper

8位的字段，包含了定义内容描述数据类型的16位二进制参数data_type的8位最高有效位，如表13-20所示。

2. marker_bit

标记位，值恒为1。

3. data_type_lower

8位字段，定义内容描述数据类型的16位二进制参数data_type的8位最低有效位，如表13-20所示。

605

4. marker_bit

标记位，值恒为1。

表13-20　MPEG-2 data_type码字

数据类型	编　码
保留	0000 0000 0000 0000
填充字节	0000 0000 0000 0001
捕获时间码	0000 0000 0000 0010
摇摄扫描参数	0000 0000 0000 0011
有效区域窗口	0000 0000 0000 0100
编码图像长度	0000 0000 0000 0101
保留	0000 0000 0000 0110
⋮	⋮
保留	1111 1111 1111 1111

5. data_length

8位的二进制数，定义了接下来的数据长度，以9位为单位表示。

注意：当data_type为“填充字节（padding_bytes）”时，下列字段将出现。这两个字段将重复多次，次数由data_length字段指定。

6. marker_bit

标记位，值恒为1。

7. padding_byte

8位字段，值为0000 0000。所有其他值都被禁止使用。

注意：当data_type为“捕获时间码（capture timecode）”时，下列字段将出现。它包含了表示帧或场的源捕获或创建时间的时间戳。这两个字段将重复多次，次数由data_length字段指定。它并不优先于任何时间码出现在系统级。

8. marker_bit

标记位，值恒为1。

9. timecode_type

2位码字，表示与图像相关的时间码数目，如表13-21所示。

表13-21 MPEG-2 timecode_type码字

时间码类型	编码
整个帧一个时间码	00
第一场或仅有的一场一个时间码	01
第二场一个时间码	10
两个时间码，两个场各一个	11

10. counting_type

可选的3位码字，它指定了用于修正nframes计数参数的方法，从而减少漂移的累积（drift accumulation）。

11. reserved_bit

保留位，值恒为0。

12. reserved_bit

保留位，值恒为0。

13. reserved_bit

保留位，值恒为0。

14. marker_bit

可选标记位，值恒为1。仅当counting_type不等于000时该字段出现。

15. nframes_conversion_code

可选位，在确定由nframes参数表示的时间量时，该位定义转换因子（1000+nframes_conversion_code）。仅当counting_type不等于000时该字段出现。

16. clock_divisor

可选的7位二进制数，定义了27MHz系统时钟的因子数，从而用于产生相同的时间戳。仅当counting_type不等于000时该字段出现。

17. marker_bit

可选标记位，值恒为1。仅当counting_type不等于000时该字段出现。

606

18. nframes_multiplier_upper

可选的8位值，为16位nframes_multiplier值的8位最高有效位。仅当counting_type不等于000时该字段出现。

19. marker_bit

可选标记位，值恒为1。仅当counting_type不等于000时该字段出现。

20. nframes_multiplier_lower

可选的8位值，为16位nframes_multiplier值的8位最低有效位。仅当counting_type不等于000时该字段出现。

下面是“场或帧捕获时间戳”信息。

21. marker_bit

可选标记位，值恒为1。仅当counting_type不等于000时该字段出现。

22. nframes

可选的8位二进制数，在导出等价时间戳时，指定了增加的帧时间增量数。仅当counting_type不等于000时该字段出现。

23. marker_bit

标记位，值恒为1。

24. time_discontinuity

1位标志位，值为1表示时间码序列中发生了中断。

25. prior_count_dropped

1位标志位，表示是否丢弃了一个或多个nframes值计数。

26. time_offset_part_a

6位值，包含了time_offset的6位最高有效位。time_offset是一个30位的符号值，定义了偏移自其他时间戳参数（当前场或帧被丢弃时，这些参数定义了相同时间戳）定义的时间的时钟周期数。

27. marker_bit

标记位，值恒为1。

28. time_offset_part_b

8位值，包含了time_offset的第二个8位最高有效位。

29. marker_bit

标记位，值恒为1。

30. time_offset_part_c

8位值，包含了time_offset的第三个8位最高有效位。

31. marker_bit

标记位，值恒为1。

32. time_offset_part_d

607　8位值，包含了time_offset的8位最低有效位。

33. marker_bit

标记位，值恒为1。

34. units_of_seconds

4位的二进制数，表示秒钟时间戳值。其值可以为0000~1001。

35. tens_of_seconds

4位的二进制数，表示10秒钟时间戳值。其值可以为0000~0101。

36. marker_bit

标记位，值恒为1。

37. units_of_minutes

4位的二进制数，表示分钟时间戳值。其值可以为0000~1001。

38. tens_of_minutes

4位的二进制数，表示10分钟时间戳值。其值可以为0000~0101。

39. marker_bit

标记位，值恒为1。

40. units_of_hours

4位的二进制数，表示小时时间戳值。其值可以为0000~1001。

41. tens_of_hours

4位的二进制数，表示10小时时间戳值。其值可以为0000~0010。

当timecode_type = 11时，"场或帧捕获时间戳"字段又再次出现，从而为第二个字段传递信息。

注意：当date_type为"摇摄参数"时，下面的字段将出现。这些字段将为有不同的高宽比的显示传输附加的摇摄信息。

42. marker_bit

标记位，值恒为1。

43. aspect_ratio_information

4为码字，与序列头所使用的高宽比信息一致。

44. reserved_bit

保留位，值恒为0。

45. reserved_bit

保留位，值恒为0。

46. reserved_bit

保留位，值恒为0。

47. display_size_present

1位标志位，表示接下来是否有display_horizontal_size和display_vertical_size字段。

48. marker_bit

标记位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

49. reserved_bit

保留位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

608

50. reserved_bit

保留位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

51. display_horizontal_size_upper

该字段为display_horizontal_size的6位最高有效位。仅当display_size_present = 1时该可选字段才出现。

52. marker_bit

标记位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

53. display_horizontal_size_lower

该字段为display_horizontal_size的8位最低有效位。仅当display_size_present = 1时该可选字段才出现。

54. marker_bit

标记位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

55. reserved_bit

保留位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

56. reserved_bit

保留位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

57. display_vertical_size_upper

该字段为display_vertical_size的6位最高有效位。仅当display_size_present = 1时该可选字段才出现。

58. marker_bit

标记位，值恒为1。仅当display_size_present = 1时该可选字段才出现。

59. display_vertical_size_lower

该字段为display_vertical_size的8位最低有效位。仅当display_size_present = 1时该可选字段才出现。

注意：对于出现的每一个帧中心偏移（frame center offsets），下面的字段都将出现。

60. marker_bit

标记位，值恒为1。

61. frame_center_horizontal_offset_upper

frame_center_horizontal_offset的8位最高有效位。frame_center_horizontal_offset的定义在图像显示扩展中给出了。

62. marker_bit

标记位，值恒为1。

63. frame_center_horizontal_offset_lower

frame_center_horizontal_offset的8位最低有效位。

64. marker_bit

标记位，值恒为1。

65. frame_center_vertical_offset_upper

frame_center_vertical_offset的8位最高有效位。frame_center_vertical_offset的定义在图像显示扩展中给出了。

66. marker_bit

标记位，值恒为1。

67. frame_center_vertical_offset_lower

609 frame_center_vertical_offset的8位最低有效位。

注意：当data_type指定为“有效区域窗口（active region window）”时，下列字段将出现。active_region_window定义了解码图像将要显示的矩形区域。该窗口不能大于由horizontal_size和vertical_size定义的区域。

68. marker_bit

标记位，值恒为1。

69. top_left_x_upper

8位的字段，是16位的top_left_x的8位最高有效位。top_left_x定义了Y采样点数目，它与top_left_y一起定义了active_region_window的左上角。

70. marker_bit

标记位，值恒为1。

71. top_left_x_lower

8位的字段，是16位的top_left_x的8位最低有效位。

72. marker_bit

标记位，值恒为1。

73. top_left_y_upper

8位的字段，是16位的top_left_y的8位最高有效位。top_left_y定义了Y扫描行数目，它与top_left_x一起定义了active_region_window的左上角。

74. marker_bit

标记位，值恒为1。

75. top_left_y_lower

8位的字段，是16位的top_left_y的8位最低有效位。

76. marker_bit

标记位，值恒为1。

77. active_region_horizontal_size_upper

8位字段，是16位的active_region_horizontal_size的8位最高有效位。active_region_horizontal_size与active_region_vertical_size一起定义了active_region_window的右下角。当active_region_horizontal_size字段值为0x000时，表示其大小未知。

78. marker_bit

标记位，值恒为1。

79. active_region_horizontal_size_lower

8位的字段，是16位的active_region_horizontal_size的8位最低有效位。

80. marker_bit

标记位，值恒为1。

81. active_region_vertical_size_upper

8位的字段，是16位的active_region_vertical_size的8位最高有效位。active_region_vertical_size与active_region_horizontal_size一起定义了active_region_window的右下角。当active_region_vertical_size字段值为0x0000时，表示其大小未知。

82. marker_bit

标记位，值恒为1。

83. active_region_vertical_size_lower

8位的字段，是16位的active_region_vertical_size的8位最低有效位。 610

注意：当data_type指定为"编码图像长度（coded picture length）"时，下面的字段将出现。

84. marker_bit

标记位，值恒为1。

85. picture_byte_count_part_a

8位字段，是32位的picture_byte_count的8位最高有效位。picture_byte_count指示了从第一个slice_start_code的第一个字节开始，到紧跟在图像最后一个宏块之后的起始码前缀之前的那个字节之间的字节数目。当值为0x0000时，表示长度是不可知的。

86. marker_bit

标记位，值恒为1。

87. picture_byte_count_part_b

8位字段，是32位的picture_byte_count的第二个8位最高有效位。

88. marker_bit

标记位，值恒为1。

89. picture_byte_count_part_c

8位字段，是32位的picture_byte_count的第三个8位最高有效位。

90. marker_bit

标记位，值恒为1。

91. picture_byte_count_part_d

8位字段，是32位的picture_byte_count的8位最低有效位。

注意：当没有其他数据出现（由data_type指示决定）时，下面两个字段将出现。这两个字段会重复多次，重复次数由data_length字段指定。

92. marker_bit

标记位，值恒为1。

93. reserved_content_description_data

保留位，长度为8位。

13.4.10 图像编码扩展

图像编码扩展只可能出现在图像头之后，其结构如图13-9所示。

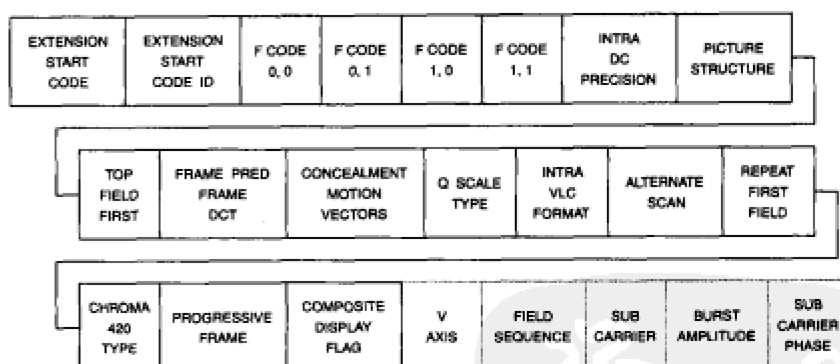


图13-9　MPEG-2图像编码扩展结构。没有给出标记位

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为1000，表示图像编码扩展的开始。

3. f_code[0,0]

4位的二进制数，取值范围为0001~1001，用于前向水平运动矢量的解码。其值不能为0000，而值为1111表示该字段将被忽略。

4. f_code[0,1]

4位的二进制数，取值范围为0001~1001，用于前向垂直运动矢量的解码。其值不能为0000，而值为1111表示该参数将被忽略。

611

5. f_code[1,0]

4位的二进制数，取值范围为0001~1001，用于后向水平运动矢量的解码。其值不能为0000，

而值为1111表示该字段将被忽略。

6. f_code[1,1]

4位的二进制数，取值范围为0001~1001，用于后向垂直运动矢量的解码。其值不能为0000，而值为1111表示该字段将被忽略。

7. intra_dc_precision

2位的码字，定义intra DC的精度，如表13-22所示。

8. picture_structure

2位的码字，定义图像结构，如表13-23所示。

表13-22 MPEG-2 intra_dc_precision码字

内部DC精度（位）	编 码
8	00
9	01
10	10
11	11

表13-23 MPEG-2 picture_structure码字

图 像 结 构	编 码
保留	00
顶场	01
底场	10
帧图像	11

612

9. top_field_first

如果progressive_sequence＝0，这一位表示解码器将最先输出哪一个场。在场中，这一位的值为0。在帧中，该位值为1表示解码帧的第一个场是顶场（top field）；该位值为0表示解码帧的第一个场是底场（bottom field）。

如果progressive_sequence＝1、repeat_first_field＝0，且该位为0，那么解码器产生一个逐行扫描帧。

如果progressive_sequence＝1、repeat_first_field＝1，且该位为0，那么解码器产生两个完全相同的逐行扫描帧。

如果progressive_sequence＝1、repeat_first_field＝1，且该位为1，那么解码器产生三个完全相同的逐行扫描帧。

10. frame_pred_frame_dct

如果这一位为1，只有frame-DCT和帧预测被采用。对于场图像，该位恒为0。如果progressive_frame为1，该参数为1。

11. concealment_motion_vectors

这一位为1，表示运动矢量是为内部宏块而编码的。

12. q_scale_type

这一位表示quantizer_scale_code和quantizer_scale之间的两个映射中的哪一个被解码器所采用。

13. intra_vlc_format

这一位表示对于内部块，DCT系数采用哪一个表。当intra_vlc_format＝0时，采用表13-36；当intra_vlc_format＝1时，采用表13-37。对于非内部块，无论intra_vlc_format值为多少，都是采用表13-36。

14. alternate_scan

这一位表示对于变换系数数据，解码器将采用哪一个扫描模式。当alternate_scan＝0时，采用图7-57所示模式；当alternate_scan＝1时，采用图7-58所示模式。

15. repeat_first_field

对于这一位的使用，请参看top_filed_first。对于场图像，这一位值为0。

16. chroma_420_type

如果chroma_format是4：2：0，那么这一位与progressive_frame是一样的。否则，这一位值为0。

17. progressive_frame

如果值为0，这一位表示帧的两个场是隔行场，两场之间有一个时间间隔。如果值为1，帧的两场来自同一时刻。

18. composite_display_flag

1位的标志位，表示v_axis、field_sequence、sub_carrier、burst_amplitude和sub_carrier_phase等是否会出现在比特流中。

19. v_axis

仅当Composite_display_flag＝1时，这一位才出现。当初始源是一个PAL视频信号时，将采用这一位。当V符号为正时，v_axis＝1；否则v_axis＝0。

这一信息可以从正在驱动MPEG-2编码器的PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将PAL编码器的V开关设置成与初始源相同。

613

20. field_sequence

仅当composite_display_flag＝1时，这3位码字才会出现。该码字定义了在最初的4场或8场序列中的场编号，如表13-24所示。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的场序列设置成与初始源相同。

表13-24 MPEG-2 field_sequence码字

帧序列	场序列	编码
1	1	000
1	2	001
2	3	010
2	4	011
3	5	100
3	6	101
4	7	110
4	8	111

21. sub_carrier

仅当composite_display_flag＝1时，这一位才出现。该位为0表示原始的副载波扫描线（subcarrier-to-line）的频率关系是正确的。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。

22. burst_amplitude

仅当composite_display_flag＝1时，这个7位的二进制数才出现。该参数定义了每个BT.601量化时（忽略MSB），初始PAL或者NTSC的脉冲幅度。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的彩色脉冲幅度设置成与初始源相同。

23. sub_carrier_phase

仅当composite_display_flag＝1时，这个8位的二进制数才出现。该参数定义了初始PAL或者NTSC副载波相位，如BT.470所定义的一样。其值定义为：(360°/256)sub_carrier_phase。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的彩色副载波相位设置成与初始源相同。

13.4.11 量化矩阵扩展

每一个量化矩阵都有默认值。当对序列头解码时，所有的矩阵复位为它们的默认值。在序列头期间或者使用这个扩展的过程中，用户定义的矩阵可能会被下载。这个可选的扩展只可能出现

在图像编码扩展之后，其结构如图13-10所示。

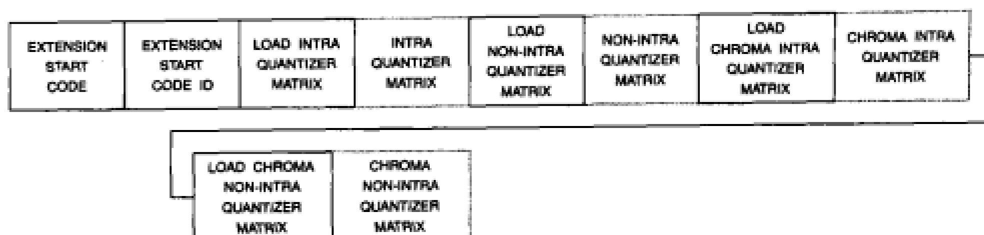


图13-10 MPEG-2量化矩阵扩展结构。没有给出标记位

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示一个新扩展数据集的开始。 614

2. extension_start_code_ID

长度为4位的字段，值为0011，表示quant_matrix_extension的开始。这个扩展同样允许量化器矩阵以4：2：2和4：4：4色度格式传送。

3. load_intra_quantizer_matrix

如果在之后跟随一个intra_quantizer_matrix，这一位设置为1。如果设置为0，那么下面的默认值将用于内部块，直到序列头或者quant_matrix_extension的下一次出现。

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

4. intra_quantizer_matrix

可选的64个8位值列表，用来替换上面给出的默认值。不允许为0。intra_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant_matrix_extension的下一次出现。出现的顺序如图7-57所示。

对于4：2：2和4：4：4数据格式，新的值同时用于Y和CbCr内部矩阵，除非装载了一个不同的CbCr内部矩阵。

5. load_non_intra_quantizer_matrix

如果其后紧跟non_intra_quantizer_matrix，此位设为1。如果设为0，采用如下默认值，直到序列头或者quant_matrix_extension的下一次出现。

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

615

6. non_intra_quantizer_matrix

可选的64个8位值列表，用来替换上面给出的默认值。不允许为0。这些数值一直有效，直到序列头或者quant_matrix_extension的下一次出现。出现的顺序如图7-57所示。

对于4：2：2和4：4：4数据格式，新的值同时用于Y和CbCr非内部矩阵，除非装载了一个新的CbCr非内部矩阵。

7. load_chroma_intra_quantizer_matrix

如果其后紧跟chroma_intra_quantizer_matrix，此位设为1。如果设为0，在使用的值中没有发生改变。如果chroma_format为4：2：0，这一位为0。

8. chroma_intra_quantizer_matrix

可选的64个8位值列表，用来替换用于CbCr数据的前一个或者默认值。不允许为0。chroma_intra_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant_matrix_extension的下一次出现。出现的顺序如图7-57所示。

9. load_chroma_non_intra_quantizer_matrix

如果其后紧跟chroma_non_intra_quantizer_matrix，此位设为1。如果设为0，在使用的值中没有发生改变。如果chroma_format为4：2：0，这一位为0。

10. chroma_non_intra_quantizer_matrix

可选的64个8位值列表，用来替换用于CbCr数据的前一个或者默认值。不允许为0。chroma_intra_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant_matrix_extension的下一次出现。出现的顺序如图7-57所示。

13.4.12　图像显示扩展

这一扩展准许显示矩阵的位置在图像之间发生移动。一个典型的应用可能就是实现摇摄和扫摄。这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-11所示。

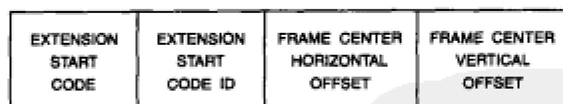


图13-11　MPEG-2图像显示扩展结构。没有给出标记位

1. extension_start_code

616

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为0111，表示图像显示扩展的开始。

在隔行扫描序列情况下，一个图像可能与一个、两个或者三个解码的场相关。因此，在比特流中有可能至少出现3个由下面4个字段组成的集。

3. frame_center_horizontal_offset

16位二进制补码，定义了水平偏移（单位为十六分之一采样点）。正数值决定解码图像的中心位置位于显示区域中心的右侧。

4. marker_bit

标记位，值恒为1。

5. frame_center_vertical_offset

16位二进制补码，定义了垂直偏移（单位为十六分之一扫描线）。正数值决定解码图像的中心位置位于显示区域中心的下方。

6. marker_bit

标记位，值恒为1。

13.4.13 图像时域可分级扩展

这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-12所示。



图13-12 MPEG-2图像时域可分级扩展结构。标记位没有给出

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为1010，表示图像时域可分级扩展的开始。

3. reference_select_code

2位的码字，表示用于预测的参考帧或场

4. forward_temporal_reference

10位的二进制数，指示用于提供前向预测的低层的时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

5. marker_bit

标记位，值恒为1。

617

6. backward_temporal_reference

10位的二进制数，指示用于提供后向预测的低层的时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

13.4.14 图像空域可分级扩展

这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-13所示。

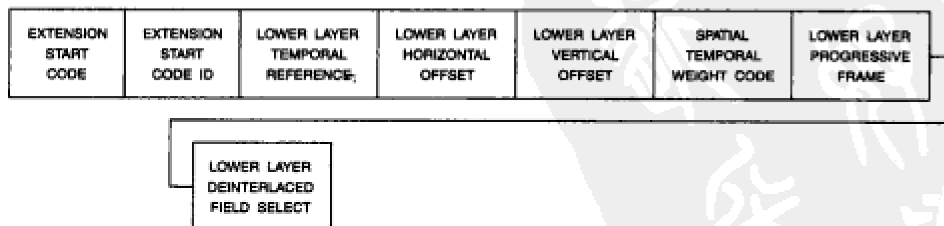


图13-13 MPEG-2图像空域可分级扩展结构。标记位没有给出

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为1001，表示图像空域可分级扩展的开始。

3. lower_layer_temporal_reference

10位的二进制数，指示用于提供预测的低层时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

4. marker_bit

标记位，值恒为1。

5. lower_layer_horizontal_offset

15位的二进制数，指示上采样低层图像的左上角相对于增强层图像的水平偏移量。对于4：2：0和4：2：2格式，这个参数必须是一个偶数。

6. marker_bit

618

标记位，值恒为1。

7. lower_layer_vertical_offset

15位的二进制数，指示上采样低层图像的左上角相对于增强层图像的垂直偏移量。对于4：2：0和4：2：2格式，这个参数必须是一个偶数。

8. spatial_temporal_weight_code_table_index

2位的码字，指示将采用哪一个空时域权重码。

9. lower_layer_progressive_frame

如果低层图像是逐行的，这一位将为1。

10. lower_layer_deinterlaced_field_select

这一位将与其他参数联合使用，以辅助解码器解码。参见表13-25。

表13-25 MPEG-2图像空域可分级扩展上采样过程

低层去隔行扫描场选择	低层逐行扫描帧	逐行扫描帧	应用去隔行扫描过程	用于预测
0	0	1	是	顶场
1	0	1	是	底场
1	1	1	否	帧
1	1	0	否	帧
1	0	0	是	两个场

13.4.15 版权扩展

这个可选的扩展只可能出现在图像编码扩展之后。

1. extension_start_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

2. extension_start_code_ID

长度为4位的字段，值为0100，表示版权扩展的开始。

3. copyright_flag

1位的标志位，如果为1，定义随后的视频内容是受版权保护的，直到下一个版权扩展的出现。如果为0，没有指示随后的视频内容是否受到版权保护。

4. copyright_identifier

8位的二进制数，指示版权拥有者。值为0000 0000表示这一信息没有提供。当copyright_flag＝0时，copyright_identifier必须为0000 0000。

5. original_or_copy

值为1表示为原始内容，为0表示为副本。

6. reserved

7位保留位，值恒为000 0000。

7. marker_bit

标记位，值恒为1。

8. copyright_number_1

表示版权编码的第44~63位的20位数。

9. marker_bit

标记位，值恒为1。

10. copyright_number_2

表示版权编码的第22~43位的22位数。

619

11. marker_bit

标记位，值恒为1。

12 copyright_number_3

表示版权编码的第0~21位的22位数。64位的copy_number唯一标识了受版权保护的内容。当copyright_identifier＝0000 0000或者copyright_flag＝0时，copyright_number必须为零。

13.4.16 摄像头参数扩展

这个可选的扩展只可能出现在图像编码扩展之后。

在值为0x000001B5的32位extension_start_code和值为1011的4位extension_start_code_ID之后，有几个字段定义所采用摄像头的位置和特性。

13.4.17 ITU-T ext. D扩展

这个可选的扩展只可能出现在图像编码扩展之后。

在值为0x000001B5的32位extension_start_code和值为1011的4位extension_start_code_ID之后，有1位数据。这一扩展的使用方法在ITU-T H.320 Annex A中给出了定义。

13.4.18 条带层

每一个条带层的数据都是由一个条带头及跟在之后的宏块数据组成的。条带层的结构在图13-5中给出了。

1. slice_start_code

32位的条带起始码，前24位值为0x000001。后8位为slice_vertical_position，其取值范围为0x01~0xAF。

slice_vertical_position定义了第一个宏块在条带中的垂直位置（以宏块为单位）。宏块的第一行的slice_vertical_position为1。

2. slice_vertical_position_extension

可选的3位二进制数，如果帧的垂直大小大于2800行，该字段表示为11位的slice_vertical_position值的3个最高有效位；否则，这个字段将不会出现。

3. priority_breakpoint

可选的7位二进制数，仅当sequence_scalable_extension出现在比特流中且scalable_mode＝“数据分区”时才出现。它定义了在比特流中的哪个位置进行分区。

4. quantizer_scale_code

5位的二进制数，取值为1~31（0值被禁用）。它定义了接收到的DCT系数的重构非零值的比例因子。解码器一直利用这个值，直到在条带层或者宏块层接收到另一个quantizer_scale_code为

止。

5. slice_extension_flag

1位可选的标志位，如果值为1，其后将出现intra_slice、slice_picture_ID_enable和slice_picture_ID字段。

6. intra_slice

可选位，仅当slice_extension_flag＝1时才出现。如果条带中的任一宏块为非内部宏块，这一位必须设置为0。

620

7. slice_picture_ID_enable

可选位，仅当slice_extension_flag＝1时才出现。值为1表示采用了slice_picture_ID。

8. slice_picture_ID

这些可选的6位字段用于辅助恢复严重的突发错误。对于一幅图像中的所有条带，slice_picture_ID值必须相同。仅当slice_extension_flag＝1时这个字段才出现。如果slice_picture_ID_enable＝0，这6位必须为00 0000

9. extra_bit_slice

1位的标志位，设置为1表示extra_information_slice跟随在其后。设置为0表示在这个字段之后没有数据。

10. extra_information_slice

如果extra_bit_slice＝1，那么这个长为9位的字段跟随在其后，其中8位为数据(extra_information_slice)，第9位为另一个extra_bit_slice，指示是否有另一个9位跟随在其后，并依此循环。

13.4.19 宏块层

每一个宏块层的数据都是由一个宏块头与跟随在其后的运动矢量和块数据组成的，如图13-5所示。

1. macroblock_escape

这个可选的11位字段是一个固定的比特字符串0000 0001 000，在当前宏块地址与前一个宏块地址之间差值大于33时采用。该字段强制macroblock_address_increment的值增加33。任意个连续的macroblock_escape字段都可能被采用。

2. macroblock_address_increment

这是一个变长的码字，定义当前宏块地址与前一个宏块地址之间的差值。它的最大值为33。当值大于33时，采用macroblock_escape字段来编码。表13-26中给出了这些变长码字。

表13-26 MPEG-2 macroblock_address_increment可变长编码表

增量值	编码	增量值	编码
1	1	9	0000 110
2	011	10	0000 1011
3	010	11	0000 1010
4	0011	12	0000 1001
5	0010	13	0000 1000
6	0001 1	14	0000 0111
7	0001 0	15	0000 0110
8	0000 111	16	0000 0101 11

（续）

增量值	编码	增量值	编码
17	0000 0101 10	25	0000 0100 000
18	0000 0101 01	26	0000 0011 111
19	0000 0101 00	27	0000 0011 110
20	0000 0100 11	28	0000 0011 101
21	0000 0100 10	29	0000 0011 100
22	0000 0100 011	30	0000 0011 011
23	0000 0100 010	31	0000 0011 010
24	0000 0100 001	32	0000 0011 001
		33	0000 0011 000
macroblock_escape			0000 0001 000

3. macroblock_type

可变长的码字，表示编码的方法和宏块的内容，如表13-27、表13-28和表13-29所示。

表13-27 MPEG-2中I图像macroblock_type的可变长编码表

类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
intra	0	0	0	0	1	0	0	1
intra，quant	1	0	0	0	1	0	0	01
空域可分级的I图像								
coded，compatible	0	0	0	1	0	0	4	1
coded，compatible，quant	1	0	0	1	0	0	4	01
intra	0	0	0	0	1	0	0	0011
intra，quant	1	0	0	0	1	0	0	0010
not coded，compatible	0	0	0	0	0	0	4	0001
SNR可分级的I图像								
coded	0	0	0	1	0	0	0	1
coded，quant	1	0	0	1	0	0	0	01
not coded	0	0	0	0	0	0	0	001

表13-28a MPEG-2中P图像macroblock_type的可变长编码表

类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
mc，coded	0	1	0	1	0	0	0	1
no mc，coded	0	0	0	1	0	0	0	01
mc，not coded	0	1	0	0	0	0	0	001
intra	0	0	0	0	1	0	0	0001 1

（续）

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
mc,coded, quant	1	1	0	1	0	0	0	0001 0
no mc,coded, quant	1	0	0	1	0	0	0	0000 1
intra,quant	1	0	0	0	1	0	0	0000 01
SNR可分级的P图像								
coded	0	0	0	1	0	0	0	1
coded, quant	1	0	0	1	0	0	0	01
not coded	0	0	0	0	0	0	0	001

表13-28b MPEG-2中P图像macroblock_type的可变长编码表

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
空域可分级的P图像								
mc，coded	0	1	0	1	0	0	0	10
mc.coded，compatible	0	1	0	1	0	1	1，2，3	011
no mc，coded	0	0	0	1	0	0	0	0000 100
no mc，coded，compatible	0	0	0	1	0	1	1，2，3	0001 11
mc，not coded	0	1	0	0	0	0	0	0010
intra	0	0	0	0	1	0	0	0000 111
mc，not coded，compatible	0	1	0	0	0	1	1，2，3	0011
mc，coded，quant	1	1	0	1	0	0	0	010
no mc,coded，quant	1	0	0	1	0	0	0	0001 00
intra,quant	1	0	0	0	1	0	0	0000 110
mc, coded, compatible, quant	1	1	0	1	0	1	1，2，3	11
no mc, coded, compatible, quant	1	0	0	1	0	1	1，2，3	0001 01
no mc, not coded, compatible	0	0	0	0	0	1	1，2，3	0001 10
coded, compatible	0	0	0	1	0	0	4	0000 101
coded, compatible, quant	1	0	0	1	0	0	4	0000 010
not coded, compatible	0	0	0	0	0	0	4	0000 0011

表13-29 MPEG-2中B图像macroblock_type的可变长编码表

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
interp, not coded	0	1	1	0	0	0	0	10
interp, coded	0	1	1	1	0	0	0	11
bwd ,not coded	0	0	1	0	0	0	0	010
bwd,coded	0	0	1	1	0	0	0	011
fwd,not coded	0	1	0	0	0	0	0	0010
fwd,coded	0	1	0	1	0	0	0	0011
intra	0	0	0	0	1	0	0	0001 1
interp,coded,quant	1	1	1	1	0	0	0	0001 0
fwd,coded, quant	1	1	0	1	0	0	0	0000 11
bwd,coded, quant	1	0	1	1	0	0	0	0000 10
intra,quant	1	0	0	0	1	0	0	0000 01
空域可分级的B图像								
interp, not coded	0	1	1	0	0	0	0	10
interp.coded	0	1	1	1	0	0	0	11
bwd ,not coded	0	0	1	0	0	0	0	010
bwd,coded	0	0	1	1	0	0	0	011
fwd,not coded	0	1	0	0	0	0	0	0010
fwd,coded	0	1	0	1	0	0	0	0011
bwd ,not coded, compatible	0	0	1	0	0	1	1,2,3	0001 10
bwd , coded, compatible	0	0	1	1	0	1	1,2,3	0001 11
fwd,not coded, compatible	0	1	0	0	0	1	1,2,3	0001 00
fwd, coded, compatible	0	1	0	1	0	1	1,2,3	0001 01
intra	0	0	0	0	1	0	0	0000 110
interp, coded, quant	1	1	1	1	0	0	0	0000 111
fwd.coded, quant	1	1	0	1	0	0	0	0000 100
bwd, coded, quant	1	0	1	1	0	0	0	0000 101
intra,quant	1	0	0	0	1	0	0	0000 0100
fwd,coded, compatible, quant	1	1	0	1	0	1	1,2,3	0000 0101
bwd,coded, compatible, quant	1	0	1	1	0	1	1,2,3	0000 0110 0
not coded, compatible	0	0	0	0	0	0	4	0000 0111 0
coded, quant, compatible	1	0	0	1	0	0	4	0000 0110 1
coded, compatible	0	0	0	1	0	0	4	0000 0111 1
SNR可分级的B图像								
coded	0	0	0	1	0	0	0	1
coded, quant	1	0	0	1	0	0	0	01
not coded	0	0	0	0	0	0	0	001

4. spatial_temporal_weight_code

可选的2位码字，表示在空域可分级情况下，时域和空域预测是如何联合为宏块进行预测的。仅当表13-27、表13-28和表13-29中的[spatial temporal weight class]＝1且spatial_temporal_weight_code_table_index≠00时该字段才出现。

5. frame_motion_type

可选的2位码字，表示如表13-30所示的宏块运动预测。仅当picture_structure＝frame，frame_pred_frame_dct＝0且表13-27、表13-28和表13-29中的[motion forward]或者[motion backward]＝1时该字段才出现。

表13-30 MPEG-2 frame_motion_type码字

空域时域权重类	预测类型	运动矢量统计	运动矢量格式	编 码
	保留			00
0,1	场	2	场	01
2,3	场	1	场	01
0,1,2,3	帧	1	帧	10
0,2,3	对偶	1	场	11

6. field_motion_type

可选的2位码字，表示如表13-31所示的宏块运动预测。仅当表13-27、表13-28和表13-29中的[motion forward]或者[motion backward]＝1且frame_motion_type字段没有出现时该字段才出现。

621

表13-31 MPEG-2 field_motion_type码字

空域时域权重类	预测类型	运动矢量统计	运动矢量格式	编 码
	保留			00
0,1	场	1	场	01
0,1	16×8 mc	2	场	10
0	对偶	1	场	11

7. dct_type

可选标志位，表示宏块是帧DCT编码还是场DCT编码。1表示是场，0表示是帧。仅当picture_structure＝11、frame_pred_frame_dct＝0并且表13-27、表13-28和表13-29中的[intra macroblock]或者[coded pattern]＝1时该标志位才出现。

8. quantizer_scale_code

可选的5位二进制数，取值为1~31（0值被禁用）。它定义了接收到的DCT系数的重构非零值的比例因子。解码器利用这个值直到接收到另一个quantizer_scale_code为止。仅当表13-27、表13-28和表13-29中的[macroblock quant]＝1时这个字段才出现。

以下为可选的运动矢量。

9. marker_bit

标记位，值恒为1。仅当concealment_motion_vectors＝1且表13-27、表13-28和表13-29中的[intra macroblock]＝1时该标记位才出现。

10. coded_block_pattern_420

这个可选的可变长码字用于推导4：2：0编码块模式（CBP），如表13-32所示。仅当表13-27、表13-28和表13-29中的[coded pattern]＝1时该字段才出现，它指示宏块中哪一个块至少有一个变换系数已被传送。编码块模式编号表示如下：

$P_1P_2P_3P_4P_5P_6$

上述编号中，对于块[n]，其任一系数出现则$P_n=1$，否则$P_n=0$。图13-2中给出了块编号（十进制格式）。

表13-32 MPEG-2 coded_block_pattern_420可变长编码表

编码块模式（CBP）	编码			编码块模式（CBP）	编码		
60	111			62	0100	0	
4	1101			24	0011	11	
8	1100			36	0011	10	
16	1011			3	0011	01	
32	1010			63	0011	00	
12	1001	1		5	0010	111	
48	1001	0		9	0010	110	
20	1000	1		17	0010	101	
40	1000	0		33	0010	100	
28	0111	1		6	0010	011	
44	0111	0		10	0010	010	
52	0110	1		18	0010	001	
56	0110	0		34	0010	000	
1	0101	1		7	0001	1111	
61	0101	0		11	0001	1110	
2	0100	1		19	0001	1101	
35	0001	1100		38	0000	1100	
13	0001	1011		29	0000	1011	
49	0001	1010		45	0000	1010	
21	0001	1001		53	0000	1001	
41	0001	1000		57	0000	1000	
14	0001	0111		30	0000	0111	
50	0001	0110		46	0000	0110	
22	0001	0101		54	0000	0101	
42	0001	0100		58	0000	0100	
15	0001	0011		31	0000	0011	1
51	0001	0010		47	0000	0011	0
23	0001	0001		55	0000	0010	1
43	0001	0000		59	0000	0010	0
25	0000	1111		27	0000	0001	1
37	0000	1110		39	0000	0001	0
26	0000	1101		0*	0000	0000	1

* 不被4：2：0色度结构所采用。

11. coded_block_pattern_1

仅当chroma_format＝4：2：2并且表13-27、表13-28和表13-29中的[coded pattern]＝1时才出现。这个可选的2位字段被用来扩展编码块模式。

12. coded_block_pattern_2

仅当chroma_format＝4：2：4并且表13-27、表13-28和表13-29中的[coded pattern]＝1时才出现。

这个可选的6位字段被用来扩展编码块模式。

13.4.20　块层

每一个块层的数据是由系数数据所组成的。图13-5给出了块层的结构。

1. dct_dc_size_luminance

仅当当前块为Y内部编码块时这个可选的可变长编码才出现，它定义了在接下来的dct_dc_differential中的位数。表13-33中给出了其取值。

表13-33　MPEG-2 dct_dc_size_luminance可变长编码表

DCT DC Size Luminance	编　码	DCT DC Size Luminance	编　码
0	100	6	1111 0
1	00	7	1111 10
2	01	8	1111 110
3	101	9	1111 1110
4	110	10	1111 1111 0
5	1110	11	1111 1111 1

2. dct_dc_differential

如果dct_dc_luminance≠0，这个可选的可变长码将出现。表13-34中给出了其取值。

表13-34　MPEG-2 dct_dc_size_chrominance可变长编码表

DCT DC Size Chrominance	编　码	DCT DC Size Chrominance	编　码
0	00	6	1111 10
1	01	7	1111 110
2	10	8	1111 1110
3	110	9	1111 11110
4	1110	10	1111 1111 10
5	11110	11	1111 1111 11

3. dct_dc_size_chrominance

仅当当前块为CbCr内部编码块时这个可选的可变长码才出现，它定义了在接下来的dct_dc_differential中的位数。表13-35中给出了其取值。

表13-35　MPEG-2 dct_dc_differential 可变长编码表

DCT DC Differential	DCT DC Size	编码（Y）	编码（CbCr）	附 加 编 码
−2048~−1024	11	111111111	1111111111	00000000000~01111111111
−1023~−512	10	111111110	1111111110	0000000000~0111111111
−511~−256	9	11111110	111111110	000000000~011111111
−255~−128	8	1111110	11111110	00000000~01111111
−127~−64	7	111110	1111110	0000000~0111111
−63~−32	6	11110	111110	000000~011111
−31~−16	5	1110	11110	00000~01111
−15~−8	4	110	1110	0000~0111
−7~−4	3	101	110	000~011
−3~−2	2	01	10	00~01

（续）

DCT DC Differential	DCT DC Size	编码（Y）	编码（CbCr）	附加编码
−1	1	00	01	0
0	0	100	00	
1	1	00	01	1
2~3	2	01	10	10~11
4~7	3	101	110	100~111
8~15	4	110	1110	1000~1111
16~31	5	1110	11110	10000~11111
32~63	6	11110	111110	100000~111111
64~127	7	111110	1111110	100000~01111111
128~255	8	1111110	11111110	10000000~11111111
256~511	9	11111110	111111110	100000000~111111111
512~1023	10	111111110	1111111110	1000000000~1111111111
1024~2047	11	111111111	1111111111	10000000000~11111111111

4. dct_dc_differential

如果dct_dc_chrominance≠0，这个可选的可变长码将出现。表13-35中给出了其取值。 622

5. dct_coefficient_first

这个可选的可变长码字用于非内部编码块中的第一个DCT系数，其定义在表13-36、表13-37、表13-38和表13-39中给出了。

表13-36 MPEG-2 dct_cofficient_first 和dct_cofficient_next的可变长编码表零

RUN	LEVEL	编码			
EOB		10			
0	1	1s	如果是第一个系数		
0	1	11s	不是第一个系数		
0	2	0100	s		
0	3	0010	1s		
0	4	0000	110s		
0	5	0010	0110	s	
0	6	0010	0001	s	
0	7	0000	0010	10s	
0	8	0000	0001	1101	s
0	9	0000	0001	1000	s
0	10	0000	0001	0011	s
0	11	0000	0001	0000	s
0	12	0000	0000	1101	0s
0	13	0000	0000	1100	1s
0	14	0000	0000	1100	0s
0	15	0000	0000	1011	1s
0	16	0000	0000	0111	11s
0	17	0000	0000	0111	10s
0	18	0000	0000	0111	01s
0	19	0000	0000	0111	00s
0	20	0000	0000	0110	11s
0	21	0000	0000	0110	10s

（续）

RUN	LEVEL	编码			
0	22	0000	0000	0110	01s
0	23	0000	0000	0110	00s
0	24	0000	0000	0101	11s
0	25	0000	0000	0101	10s
0	26	0000	0000	0101	01s
0	27	0000	0000	0101	00s
0	28	0000	0000	0100	11s
0	29	0000	0000	0100	10s
0	30	0000	0000	0100	01s
0	31	0000	0000	0100	00s
0	32	0000	0000	0011	000s
0	33	0000	0000	0010	111s
0	34	0000	0000	0010	110s
0	35	0000	0000	0010	101s
0	36	0000	0000	0010	100s
0	37	0000	0000	0010	011s
0	38	0000	0000	0010	010s
0	39	0000	0000	0010	001s
0	40	0000	0000	0010	000s
1	1	011s			
1	2	0001	10s		
1	3	0010	0101	s	
1	4	0000	0011	00s	
1	5	0000	0001	1011	s
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
1	8	0000	0000	0011	111s
1	9	0000	0000	0011	110s
1	10	0000	0000	0011	101s
1	11	0000	0000	0011	100s
1	12	0000	0000	0011	011s
1	13	0000	0000	0011	010s
1	14	0000	0000	0011	001s
1	15	0000	0000	0001	0011s
1	16	0000	0000	0001	0010s
1	17	0000	0000	0001	0001s
1	18	0000	0000	0001	0000s
2	1	0101	s		
2	2	0000	100s		
2	3	0000	0010	11s	
2	4	0000	0001	0100	s
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0100	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0011	0s		

（续）

RUN	LEVEL	编	码		
4	2	0000	0011	11s	
4	3	0000	0001	0010	s
5	1	0011	11s		
5	2	0000	0010	01s	
5	3	0000	0000	1001	0s
6	1	0011	01s		
6	2	0000	0001	1110	s
6	3	0000	0000	0001	0100s
7	1	0001	00s		
7	2	0000	0001	0101	s
8	1	0000	111s		
8	2	0000	0001	0001	s
9	1	0000	101s		
9	2	0000	0000	1000	1s
10	1	0010	0111	s	
10	2	0000	0000	1000	0s
11	1	0010	0011	s	
11	2	0000	0000	0001	1010s
12	1	0010	0010	s	
12	2	0000	0000	0001	1001s
13	1	0010	0000	s	
13	2	0000	0000	0001	1000s
14	1	0000	0011	10s	
14	2	0000	0000	0001	0111s
15	1	0000	0011	01s	
15	2	0000	0000	0001	0110s
16	1	0000	0010	00s	
16	2	0000	0000	0001	0101s
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
27	1	0000	0000	0001	1111s
28	1	0000	0000	0001	1110s
29	1	0000	0000	0001	1101s
30	1	0000	0000	0001	1100s
31	1	0000	0000	0001	1011s
ESC		0000	01		

注：s表示非零系数（level）的符号，0表示正，1表示负。

表13-37　MPEG-2 dct_cofficient_first和dct_cofficient_next的可变长编码表一

RUN	LEVEL	编码			
EOB		0110			
0	1	10s			
0	2	110s			
0	3	0111	s		
0	4	1110	0s		
0	5	1110	1s		
0	6	0001	01s		
0	7	0001	00s		
0	8	1111	011s		
0	9	1111	100s		
0	10	0010	0011	s	
0	11	0010	0010	s	
0	12	1111	1010	s	
0	13	1111	1011	s	
0	14	1111	1110	s	
0	15	1111	1111	s	
0	16	0000	0000	0111	11s
0	17	0000	0000	0111	10s
0	18	0000	0000	0111	01s
0	19	0000	0000	0111	00s
0	20	0000	0000	0110	11s
0	21	0000	0000	0110	10s
0	22	0000	0000	0110	01s
0	23	0000	0000	0110	00s
0	24	0000	0000	0101	11s
0	25	0000	0000	0101	10s
0	26	0000	0000	0101	01s
0	27	0000	0000	0101	00s
0	28	0000	0000	0100	11s
0	29	0000	0000	0100	10s
0	30	0000	0000	0100	01s
0	31	0000	0000	0100	00s
0	32	0000	0000	0011	000s
0	33	0000	0000	0010	111s
0	34	0000	0000	0010	110s
0	35	0000	0000	0010	101s
0	36	0000	0000	0010	100s
0	37	0000	0000	0010	011s
0	38	0000	0000	0010	010s
0	39	0000	0000	0010	001s
0	40	0000	0000	0010	000s
1	1	010s			
1	2	0011	0s		

（续）

RUN	LEVEL	编码			
1	3	1111	001s		
1	4	0010	0111	s	
1	5	0010	0000	s	
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
1	8	0000	0000	0011	111s
1	9	0000	0000	0011	110s
1	10	0000	0000	0011	101s
1	11	0000	0000	0011	100s
1	12	0000	0000	0011	011s
1	13	0000	0000	0011	010s
1	14	0000	0000	0011	001s
1	15	0000	0000	0001	0011s
1	16	0000	0000	0001	0010s
1	17	0000	0000	0001	0001s
1	18	0000	0000	0001	0000s
2	1	0010	1s		
2	2	0000	111s		
2	3	1111	1100	s	
2	4	0000	0011	00s	
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0110	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0001	10s		
4	2	1111	1101	s	
4	3	0000	0001	0010	s
5	1	0001	11s		
5	2	0000	0010	0s	
5	3	0000	0000	1001	0s
6	1	0000	110s		
6	2	0000	0001	1110	s
6	3	0000	0000	0001	0100s
7	1	0000	100s		
7	2	0000	0001	0101	s
8	1	0000	101s		
8	2	0000	0001	0001	s
9	1	1111	000s		
9	2	0000	0000	1000	1s
10	1	1111	010s		
10	2	0000	0000	1000	0s
11	1	0010	0001	s	

（续）

RUN	LEVEL	编	码		
11	2	0000	0000	0001	1010s
12	1	0010	0101	s	
12	2	0000	0000	0001	1001s
13	1	0010	0100	s	
13	2	0000	0000	0001	1000s
14	1	0000	0010	1s	
14	2	0000	0000	0001	0111s
15	1	0000	0011	1s	
15	2	0000	0000	0001	0110s
16	1	0000	0011	01s	
16	2	0000	0000	0001	0101s
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
27	1	0000	0000	0001	1111s
28	1	0000	0000	0001	1110s
29	1	0000	0000	0001	1101s
30	1	0000	0000	0001	1100s
31	1	0000	0000	0001	1011s
ESC		0000	01		

注：s表示非零系数（level）的符号，0表示正，1表示负。

表13-38　dct_cofficient_first和dct_cofficient_next的游程编码及其后面的转义码

游程编码	编	码
0	0000	00
1	0000	01
2	0000	10
⋮	⋮	⋮
62	1111	10
63	1111	11

表13-39　dct_cofficient_first和dct_cofficient_next的LEVEL编码及其后面的转义码

LEVEL	编	码	
−2047	1000	0000	0001
−2046	1000	0000	0010
		⋮	
−1	1111	1111	1111
禁止	0000	0000	0000
1	0000	0000	0001
		⋮	
2047	0111	1111	1111

6. dct_coefficient_next

最多63个可选的可变长码字，仅对I、P和B帧出现。它们是第一个DCT系数之后的其他DCT系数，其定义在表13-36、表13-37、表13-38和表13-39中给出了。

7. end_of_block

2位或者4位值，用于指示没有其他非零系数出现。这个参数的值为10或者0110。

13.5 运动补偿

图13-14说明了基本的运动补偿过程。运动补偿形成了从前面的解码图像得到的预测，这些预测结果依次与来自IDCT的系数数据（误差项）组合。

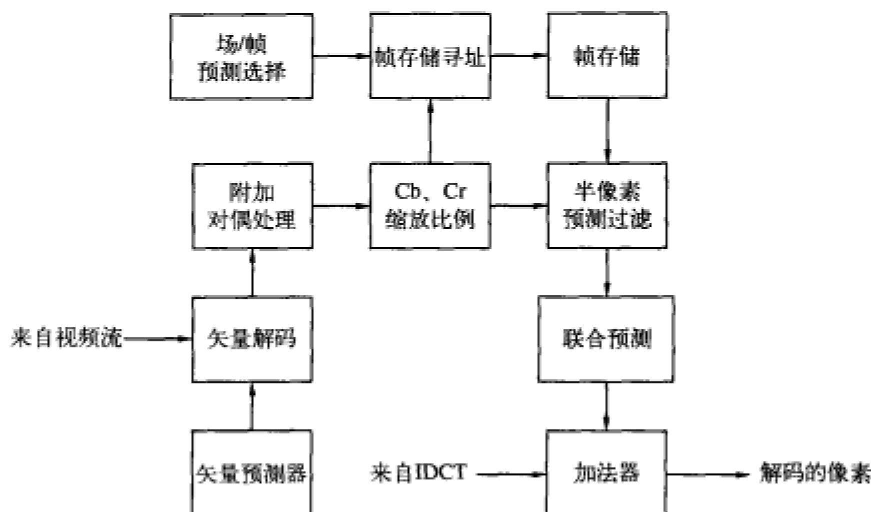


图13-14 简化的运动补偿过程

13.5.1 场预测

P图像的预测是由两个最近解码的参考场所构成的。图13-15给出了最简单的情况，当预测一帧的第一个图像或者在一帧内采用场预测时使用。

预测一帧的第二个场同样需要两个最近解码的参考场。图13-16和图13-17给出了这种情况的例子，两图中给出的第二个图像分别是底场和顶场。

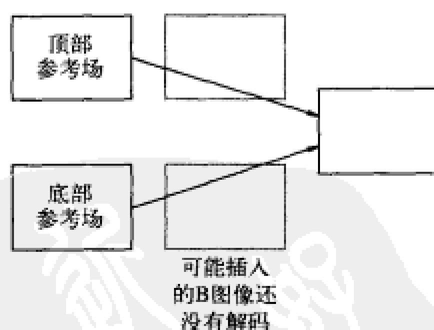


图13-15 第一场的P图像预测或者帧图像中的场预测

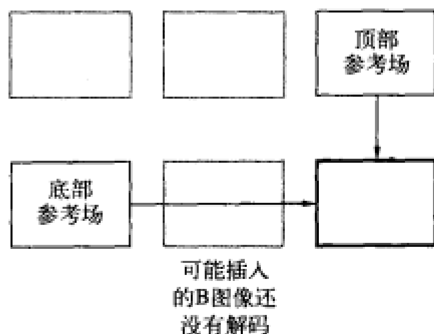


图13-16 第二场图像的P图像预测（底场）

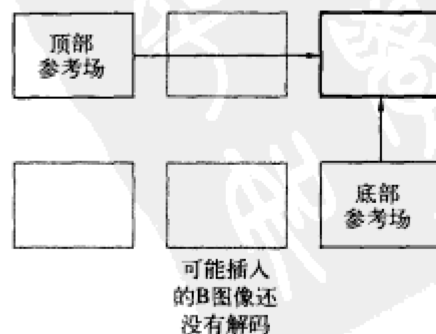


图13-17 第二场图像的P图像预测（顶场）

B图像的场预测是由两个最近的参考帧的两个场所构成的，如图13-18所示。

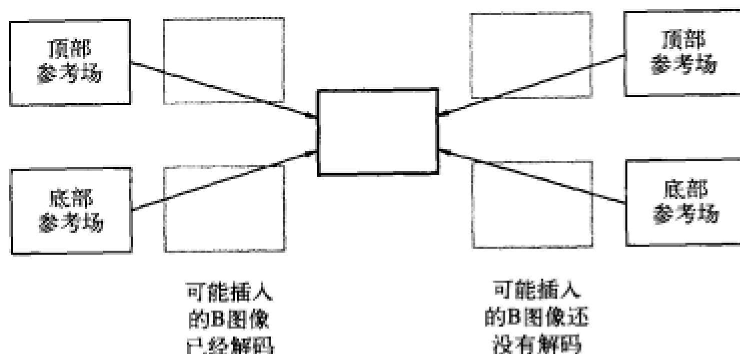


图13-18 B场或者帧图像的场预测

13.5.2 帧预测

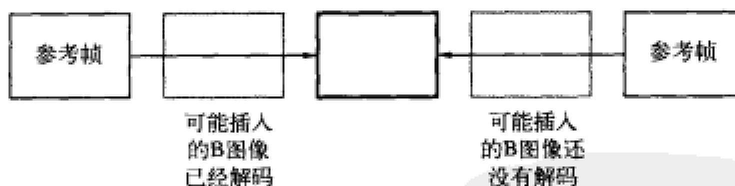
P图像的预测是由两个最近的解码图像所构成的，如图13-19所示。参考的图像可能已经作为两个场或者一个单独的帧而解码了。



图13-19 P图像的帧预测

B图像的预测是由两个最近的参考帧所构成的，如图13-20所示。每一个参考帧可能已经作为两个场或者一个单独的帧而解码了。

623~642



643~646

图13-20 B图像的帧预测

13.6 PES包

一个打包基本流（PES）由一个单独的基本流组成，该基本流已经被封装成数据包，每个数据包由一个附加的包头开始。一个PES包含来自一个源的唯一类型的数据（音频、视频等）。

图13-21给出了PES包的一般格式。注意，起始码（0x000001xx）必须是字节对齐，其实现方式是通过在起始码之前插入0~7个0。

1. packet_start_code_prefix

长度为24位的字段，值为0x000001，与stream_ID一起指示一个包的开始。

2. stream_ID

长度为8位的码，定义了基本流的类型和数目，如表13-40所示。对于ATSC和OpenCable™标准，音频流的值必须是1011 1101，用来指示Dolby® Digital。

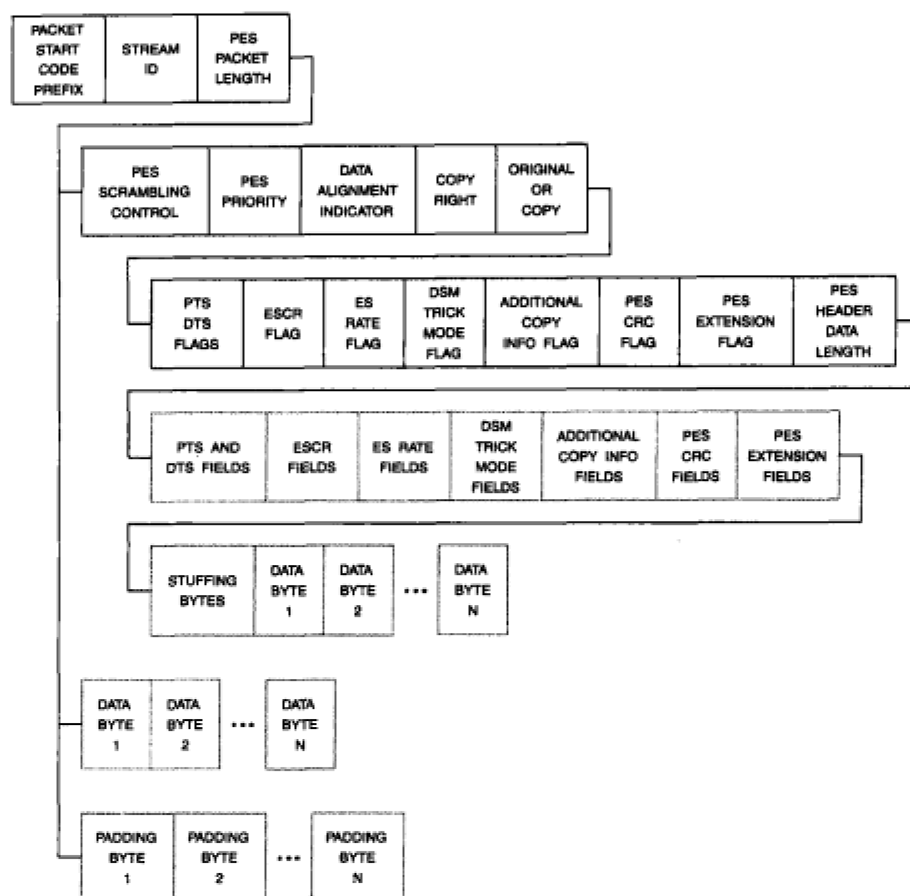


图13-21 MPEG-2 PES包结构。标记位和保留位没有给出

表13-40 MPEG-2 stream_ID码字

流	编 码
所有音频流	1011 1000
所有视频流	1011 1001
节目流映射	1011 1100
私有流1（包括LPCM、Dolby Digital、Dolby Digital Plus、DTS、DTS-HD和MLP audio）	1011 1101
填充流	1011 1110
私有流2	1011 1111
MPEG-1.3、-2.3、-4.3或者MPEG-2.7音频流	110x xxxx
MPEG-1.2、-2.2、-4.2或者MPEG-4.10（H.264）视频流	1110 xxxx
ECM流	1111 0000
EMM流	1111 0001
DSM-CC流	1111 0010
ISO/IEC 13552流	1111 0011
ITU-T H.222.1类型A	1111 0100
ITU-T H.222.1类型B	1111 0101

（续）

流	编 码
ITU-T H.222.1类型C	1111 0110
ITU-T H.222.1类型D	1111 0111
ITU-T H.222.1类型E	1111 1000
辅助流	1111 1001
MPEG-4 SL打包流	1111 1010
MPEG-4 FlexMux流	1111 1011
元数据流	1111 1100
扩展流ID	1111 1101
保留	1111 1110
节目流目录PSD	1111 1111

3. PES_packet_length

16位的二进制数，定义了在这个字段之后PES包中的字节数。值为零表示既没有定义也不是有界的，且零值仅用于传输流。对于ATSC标准的视频流，其值必须是0x0000。

注意：如果stream_ID＝program stream map、padding stream、private stream 2、ECM stream、EMM stream、DSM-CC stream、H.222.1 type E或者program stream directory，下面的字段（直到下一个节点）将不会出现。

4. marker_bits

可选的2位标记位，值为10。

5. PES_scrambling_control

可选的两位编码，定义了加扰模式（scrambling mode）。00表示不加扰，01表示保留，10表示每一个键都加扰，11表示奇数键加扰。对于SVCD、ATSC和OpenCable™标准，其值必须为00。

6. PES_priority

可选位，定义了PES包载荷的优先级。1的优先级比0高。对于DVB标准，这个字段是可选的，如果出现了，解码器可以忽略掉。对于SVCD标准，其值必须为0。

7. data_alignment_indicator

可选位，值为1表示PES包头之后紧跟视频起始码，或者由数据流对齐描述符（如果有）定义的音频同步字。对于SVCD标准，其值必须为0。

647

表13-41 某些常用的MPEG-2 stream_ID_extension码字

流	编 码
IPMP控制信息流	000 0000
IPMP流	000 0001
保留数据流	000 0010
⋮	
保留数据流	011 1111
私有流	100 0000
⋮	
私有流	101 0100
SMPTE421M（VC-1）视频流	101 0101
Dolby Digital、Dolby Digital Plus、DTS或DTS-HD核心音频流	111 0001
MLP或DTS-HD扩展音频流	111 0010

8. copyright

标志位，1表示素材是有版权的。对于SVCD，这个值必须为0。

9. original_or_copy

标志位，1表示该素材是原件；0表示该素材是一份副本。对于SVCD标准，这个值必须为1。

10. PTS_DTS_flags

两位标志位，值10表示一个PTS（显示时间戳）字段将出现在PES包头中。值11表示PTS和DTS（解码时间戳）字段将同时出现。值00表示PTS和DTS字段都不出现。对于SVCD标准，这个值必须为00、10或者11（11仅对视频）。

648~649

11. ESCR_flag

标志位，1表示ESCR（基本流时钟参考）基准和扩展字段将出现在PES包头中。对于SVCD、ATSC和OpenCable™标准，ESCR字段是可选的，如果出现了，解码器可以将其忽略。

12. ES_rate_flag

标志位，1表示ES_rate（基本流速率）将出现在PES包头中。对于SVCD、ATSC和OpenCable™标准，这个值必须为0。对于DVB标准，ES字段是可选的，如果出现了，解码器可以将其忽略。

13. DSM_trick_mode_flag

标志位，1表示trick_mode_flag字段将出现。对于SVCD标准，这个值必须为0。

14. additional_copy_info_flag

标志位，1表示additional_copy_info字段将出现。对于SVCD标准，这个值必须为0。

15. PES_CRC_flag

标志位，1表示previous_PES_packet_CRC字段将出现。对于SVCD、ATSC和OpenCable™标准，这个值必须为0。

16. PES_extension_flag

标志位，1表示一个扩展字段将出现在PES包头中。当传递SMPTE 421M（VC-1）视频流时，这一位必须为1，从而在PES包头中插入扩展。

17. PES_header_data_length

8位的二进制数，定义了PES包头中的可选字段和填充位的字节数。

18. marker_bits

可选的4位标记位，值为0010。仅当PTS_DTS_flag＝10时这个字段才出现。

19. PTS [32~30]

可选字段，仅当PTS_DTS_flag＝10时才出现。

20. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝10时这个字段才出现。

21. PTS [29~15]

可选字段，仅当PTS_DTS_flag＝10时才出现。

22. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝10时这个字段才出现。

23. PTS [14~0]

可选的33位显示时间戳（PTS），指示解码器计划的显示时间。它所采用的单位是27MHz时钟除以300。仅当PTS_DTS_flag＝10时这个字段才出现。

24. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝10时这个字段才出现。

650

25. marker_bits

可选的4位标记位，值为0011。仅当PTS_DTS_flag＝11时这个字段才出现。

26. PTS [32~30]

可选字段，仅当PTS_DTS_flag＝11时才出现。

27. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝11时这个字段才出现。

28. PTS [29~15]

可选字段，仅当PTS_DTS_flag＝11时才出现。

29. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝11时这个字段才出现。

30. PTS [14~0]

可选字段，仅当PTS_DTS_flag＝11时才出现。

31. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝10时这个字段才出现。

32. marker_bits

可选的4位标记位，值为0001。仅当PTS_DTS_flag＝11时这个字段才出现。

33. DTS [32~30]

可选字段，仅当PTS_DTS_flag＝11时才出现。

34. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝11时这个字段才出现。

35. DTS [29~15]

可选字段，仅当PTS_DTS_flag＝11时才出现。

36. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝11时这个字段才出现。

37. DTS [14~0]

可选的33位解码时间戳（DTS），指示计划的解码时间。它所采用的单位是27MHz时钟除以300。仅当PTS_DTS_flag＝11时这个字段才出现。

38. marker_bit

可选的标记位，值为1。仅当PTS_DTS_flag＝10时这个字段才出现。

39. reserved_bits

可选的两位保留位，值为11。仅当ESCR_flag＝1时这个字段才出现。

40. ESCR_base [32~30]

可选字段，仅当ESCR_flag＝1时才出现。

41. marker_bit

可选的标记位，值为1。仅当ESCR_flag＝1时这个字段才出现。

42. ESCR_base [29~15]

可选字段，仅当ESCR_flag＝1时才出现。

43. marker_bit

651

可选的标记位，值为1。仅当ESCR_flag＝1时这个字段才出现。

44. ESCR_base [14~0]

可选字段，仅当ESCR_flag＝1时才出现。

45. marker_bit

可选的标记位，值为1。仅当ESCR_flag＝1时这个字段才出现。

46. ESCR_extension

可选的9位基本流时钟参考（ESCR）扩展和33位ESCR基本信息组合成一个42位的值。它表示包含ESCR_base最后一位的字节计划到达的时间。ESCR_base的值定义了90kHz时钟周期的数目。

ESCR_extension的值定义了在90kHz时钟周期启动之后27MHz时钟周期的数目。

这个字段仅当ESCR_flag＝1时才出现。

47. marker_bit

可选的标记位，值为1。仅当ESCR_flag＝1时这个字段才出现。

48. marker_bit

可选的标记位，值为1。仅当ESCR_rate_flag＝1时这个字段才出现。

49. ES_rate

可选的22位基本流速率（ES_rate），指示解码器接收PES包字节的速率。其定义的单位为50字节/秒。仅当ESCR_rate_flag＝1时这个字段才出现。

50. marker_bit

可选的标记位，值为1。仅当ESCR_rate_flag＝1时这个字段才出现。

51. trick_mode_control

这个可选的3位码字指示哪一个窍门模式（trick mode）被应用到了视频流，如表13-42所示。仅当DSM_trick_mode_flag＝1时这个字段才出现。

表13-42 MPEG-2 trick_mode_control码字

窍门模式	编码	窍门模式	编码
快进	000	慢退	100
慢进	001	保留	101
冻结帧	010	保留	110
快退	011	保留	111

52. field_ID

可选的2位码字，指示哪一个场将被显示，如表13-43所示。仅当DSM_trick_mode_flag＝1且trick_mode_control＝000或者011时这个字段才出现。

53. intra_slice_refresh

可选的1位标志位，1表示条带之间可能有宏块丢失。仅当DSM_trick_mode_flag＝1且trick_mode_control＝000或者011时这个字段才出现。

652

54. frequency_truncation

可选的2位码字，表示一个受限的系数集可能被用于数据编码中，如表13-44所示。仅当DSM_trick_mode_flag＝1且trick_mode_control＝000或者011时这个字段才出现。

表13-43 MPEG-2 field_ID码字

场ID	编码
仅顶场	00
仅底场	01
完全帧	10
保留	11

表13-44 MPEG-2 frequency_truncation码字

描述	编码
仅DC系数是非零值	00
前3个系数是非零值	01
前6个系数是非零值	10
所有系数都可能是非零值	11

55. rep_cntrl

可选的5位二进制数，表示每一个隔行场或者逐行帧应该显示的次数。值00000被禁用。仅当DSM_trick_mode_flag＝1且trick_mode_control＝001或者100时这个字段才出现。

56. field_ID

表13-42中给出的可选的2位码字，表示将要显示哪一个场。仅当DSM_trick_mode_flag＝1且trick_mode_control＝010时这个字段才出现。

57. reserved_bits

可选的3位保留位，值为111。仅当DSM_trick_mode_flag＝1且trick_mode_control＝010时这个字段才出现。

58. reserved_bits

可选的5位保留位，值为1 1111。仅当DSM_trick_mode_flag＝1且trick_mode_control＝101、110或者111时这个字段才出现。

59. marker_bit

可选的标记位，值恒为1。仅当additional_copy_info_flag＝1时这个字段才出现。

60. additional_copy_info

可选的7位字段，包含关于版权信息的私有数据。仅当additional_copy_info_flag＝1时这个字段才出现。

61. previous_PES_packet_CRC

可选的16位校验位，仅当PES_CRC_flag＝1时出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

62. PES_private_data_flag

可选的1位标志位，值为1时表示私有数据将出现。仅当PES_extension_flag＝1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。

63. packet_header_field_flag

1位标志位，值为1时表示一个MPEG-1包头或者节目流包头出现在PES包头中。仅当PES_extension_flag＝1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。

653

64. program_packet_sequence_counter_flag

可选的1位标志位，值为1时表示program_packet_sequence_counter、MPEG1_MPEG2_identifier和original_stuff_length等字段将出现。仅当PES_extension_flag＝1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。

65. P-STD_buffer_flag

可选的1位标志位，值为1时表示P-STD_buffer_scale和P-STD_buffer_size将出现。仅当PES_extension_flag＝1时这个字段才出现。对于ATSC和OpenCable™标准，其值必须为0。对于SVCD标准，其值必须为1。

66. reserved_bits

可选的3位保留位，值恒为111。仅当PES_extension_flag＝1时这个字段才出现。

67. PES_extension_flag_2

可选的1位标志位，值为1时表示PES_extension_field_length及其相关字段将出现。仅当PES_extension_flag＝1时这个字段才出现。对于SVCD标准，其值必须为0。当传输SMPTE 421M（VC-1）视频流时，这一位必须为1，从而使得能够将第二个扩展组插入到PES包头中。

68. PES_private_data

可选的128位私有数据，与其前后的字段组合在一起，不能与packet_start_code_prefix相似。仅当PES_extension_flag = 1并且PES_private_data_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

69. pack_field_length

可选的8位二进制数，表示紧跟在包组头之后的字段的长度，单位是字节。仅当PES_extension_flag = 1并且pack_header_field_flag = 1时这个字段以及紧跟在包组头之后的字段才出现。

70. marker_bit

可选的标记位，值恒为1。仅当PES_extension_flag = 1并且program_packet_sequence_counter_flag = 1时这个字段才出现。

71. program_packet_sequence_counter

可选的7位二进制数，对于节目流或者MPEG-1系统流中每一个连续的PES包，其值增加1。当达到最大值后，其值反转到0。对于两个连续的PES包，其值不能相同。仅当PES_extension_flag = 1并且program_packet_sequence_counter_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

72. marker_bit

可选的标记位，值恒为1。仅当PES_extension_flag = 1并且program_packet_sequence_counter_flag = 1时这个字段才出现。

73. MPEG1_MPEG2_identifier

可选标志位，值为1时表示PES包有来自MPEG-1系统流的信息。值为0表示PES包有来自节目流的信息。仅当PES_extension_flag = 1并且program_packet_sequence_counter_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

74. original_stuff_length

可选的6位二进制数，指定了用于初始PES或者MPEG-1包头中填充的字节数。仅当PES_extension_flag = 1并且program_packet_sequence_counter_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。 654

75. marker_bits

可选的2位标记位，值恒为01。仅当PES_extension_flag = 1并且P-STD_buffer_flag = 1时这个字段才出现。

76. P-STD_buffer_scale

可选位，表示接下来的P-STD_buffer_size参数的比例因子。对于音频流，其值为0。对于视频流，其值为1。对于所有其他类型的流，其值既可为0，也可为1。仅当PES_extension_flag = 1并且P-STD_buffer_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

77. P-STD_buffer_size

可选的13位二进制数，指定了解码器输入缓冲区的大小。如果P-STD_buffer_scale值为0，该字段为128字节。如果P-STD_buffer_scale值为1，该字段为1024字节。仅当PES_extension_flag = 1并且P-STD_buffer_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

78. marker_bit

可选的标记位，值恒为1。仅当PES_extension_flag = 1并且PES_extension_flag_2 = 1时这个字段

才出现。

79. PES_extension_field_length

可选的7位二进制数，表示下面三个字段总的字节数。仅当PES_extension_flag＝1并且PES_extension_flag_2＝1时这个字段才出现。

80. Stream_ID_extension_flag

可选的1位标志位，值为0时表示接下来是一个stream_ID_extension字段。仅当PES_extension_flag＝1并且PES_extension_flag_2＝1时这个字段才出现。当传输SMPTE 421M（VC-1）视频流时，这一位必须为0，从而使得能够插入一个有效的stream_id_extension。

81. stream_ID_extension

7位码字，作为stream_ID的扩展使用，用来指定表13-41中所定义的基本流类型。仅当stream_ID＝1111 1101时这个字段才被采用。仅当PES_extension_flag和PES_extension_flag_2＝1并且stream_ID_extension_flag＝0时这个可选的字段才出现。

82. reserved_byte

[n]字节保留数据，值为1111 1111。仅当PES_extension_flag和PES_extension_flag_2＝1并且stream_ID_extension_flag＝0时这个可选的字段才出现。

83. stuffing_byte

可选的[n]字节填充数据，值为1111 1111。最多可以采用32字节填充数据。这些数据将被解码器忽略。

655

84. PES_packet_data_byte

来自音频流、视频流、私有流1、辅助流、H.222.1类型A-D流或者ISO/IEC 13552流的[n]字节数据。字节数由PES_packet_length字段推导而得到。

注意：当stream_ID＝节目流映射、私有流2、ECM流、EMM流、DSM-CC流、H.2221类型E或者节目流目录时，下面的字段将出现。

85. PES_packet_data_byte

来自节目流映射、私有流2、ECM流、EMM流、DSM-CC流、H.2221类型E流或者节目流目录描述符的[n]字节数据。字节数由PES_packet_length字段推导而得到。

注意：当stream_ID＝padding stream时，下面的字段将出现。

86. padding_byte

[n]字节，值为1111 1111。字节数由PES_packet_length字段指定。这一字段将被解码器忽略。

13.7　节目流

节目流是设计用于相对无差错的环境中的，它被DVD和SVCD标准所采用。节目流由一个或多个PES包复用到一起所组成，并且与数据一起编码，从而使得它们能够同时解码。节目流包可以是可变长的，且可以相对非常长。

图13-22给出了节目流的一般格式。注意，起始码（0x000001xx）必须是字节对齐的，可以通过在起始码之前插入0~7个0位来实现。

656

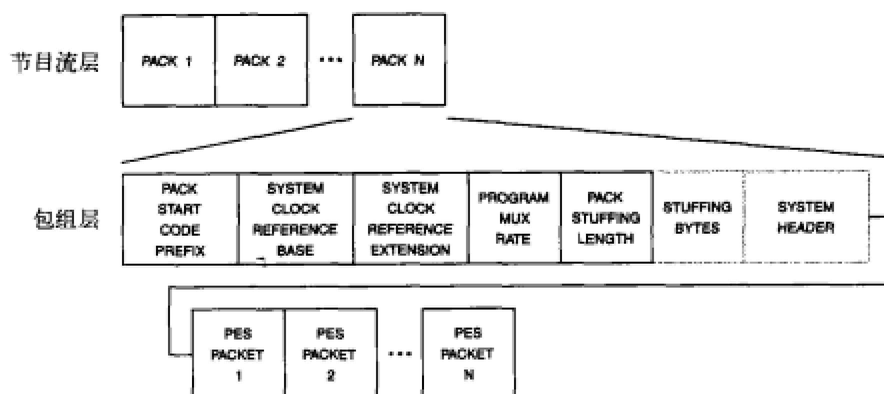


图13-22 MPEG-2节目流结构。没有给出标记和保留位

13.7.1 包组层

每个包组的数据由一个包组头跟一个可选的系统头和一个或多个PES包组成。

1. pack_start_code

32位字符串，值为0x000001BA，指示一个包组的开始。

2. marker_bits

2位标记位，值为01。

3. system_clock_reference_base [32~30]

4. marker_bit

标记位，值为1。

5. system_clock_reference_base [29~15]

6. marker_bit

标记位，值为1。

7. system_clock_reference_base [14~0]

8. marker_bit

标记位，值为1。

9. system_clock_reference_extension

9位字段，与system_clock_reference_base字段一起组成系统时钟参考（SCR）。system_clock_reference_base字段采用的单位是90kHz乘以1/300。system_clock_reference_extension字段采用单位是27MHz。SCR指示包含system_clock_reference_base字段最后一位的字节的预期达到时间。

10. marker_bit

标记位，值为1。

11. program_mux_rate

22位二进制数，定义了一个以每秒50字节为度量的值，该值禁止取零。它指示解码器接收节目流的速度。对于SVCD标准，该值必须小于等于6972_D。

12. marker_bit

标记位，值为1。

13. marker_bit

标记位，值为1。

14. reserved_bits

5位保留位，值为1 1111。

15. pack_stuffing_length

3位二进制数，定义了这一字段之后stuffing_byte字段的个数。

16. stuffing_byte

可能出现0~7个填充字节。解码器忽略这些字节。每个字节的值为1111 1111。

13.7.2 系统头

该字段包含比特流参数的梗概。在第一个包组头后必然会跟随一个系统头，但在未来的包组头中它可能会可选地重复出现。

1. system_header_start_code

657

32位字符串，值为0x000001BB，指示一个系统头的开始。

2. header_length

16位二进制数，定义了这个字段之后系统头的字节数。

3. marker_bit

标记位，值为1。

4. rate_bound

22位二进制数，定义了一个大于或等于program_mux_rate最大值的值。对于SVCD标准，该值必须为6972_D。

5. marker_bit

标记位，值为1。

6. audio_bound

6位二进制数，定义了一个大于或等于有效音频流最大个数的值（0~32）。对于SVCD标准，该值必须为0。

7. fixed_flag

1位标志位，值为1时表示固定的位速操作。值为0时表示可变的位速操作。对于SVCD标准，该值必须为0。

8. CSPS_flag

1位标志位。值为1表示节目流是“受约束的系统参数流”。

9. system_audio_lock_flag

1位标志位。如果在音频采样速率和解码器系统时钟频率之间存在已定义的、稳定的关系，该字段值为1。对于SVCD标准，该值必须为1。

10. system_video_lock_flag

1位标志位。如果在视频图像速率和解码器系统时钟频率之间存在已定义的、稳定的关系，该位值为1。对于SVCD标准，该值必须为1。

11. marker_bit

标记位，值为1。

12. video_bound

5位二进制数，定义了一个大于或等于有效视频流最大个数的值（0~16）。对于SVCD标准，该值必须为0或1。

13. reserved_bits

7位保留位，值为111 1111。

14. stream_ID

可选的8位编码，如表13-40所示，这一字段指示应用于PSTD_buffer_bound_scale 和PSTD_buffer_size_bound字段的流。

15. marker_bits

可选的2位标记位，值为11。仅当stream_ID出现时这个字段才出现。

16. P-STD_buffer_bound_scale

仅当stream_ID出现时这个可选位才出现，指示用于P-STD_buffer_size_bound的缩放比例因子。值为0表示stream_ID定义一个音频流。值为1表示stream_ID定义一个视频流。对于其他类型的流ID，该值要么为0，要么为1。

658

17. P-STD_buffer_size_bound

可选的13位二进制数，定义了一个大于或等于解码器输入缓冲最大容量的值。仅当stream_ID出现时这个字段才出现。如果P-STD_buffer_bound_scale的值为0，该字段包含128字节。如果P-STD_buffer_bound_scale的值为1，该字段包含1024字节。

13.7.3 节目流映射

节目流映射（Program Stream Map，PSM）提供节目流中的比特流以及比特流相互之间关系的描述。如果stream_ID＝program stream map，那么PSM以PES包数据的形式出现。

1. packet_start_code_prefix

24位字符串，值为0x000001，指示一个流映射的开始。

2. map_stream_ID

8位字符串，值为1011 1100。

3. program_stream_map_length

16位二进制数，指示这个字段之后的字节数。最大值为1018_D。

4. current_next_indicator

标志位。值为1表示节目流映射目前可用，值为0表示节目流映射还不可用，但是下一个将有效。

5. reserved_bits

2位保留位，值为11。

6. program_stream_map_version

5位二进制数，定义了节目流映射的版本号。当节目流映射变化时这一字段必须加一，且其值一旦达到31就重新归零。

7. reserved_bits

7位保留位，值为111 1111。

8. marker_bit

标记位，值为1。

9. program_stream_info_length

16位二进制数，定义了紧跟在该字段之后的描述符总长（单位是字节）。

10. descriptor_loop

deescriptor_loop中可能出现不同类型的描述符。

11. elementary_stream_map_length

16位二进制数，指示节目流映射中所有基本流信息的字节数。

注意：对于每个具有唯一stream_type值的流，下列4个字段出现。

12. stream_type

8位码字，定义了表13-45给出的流类型。

13. elementary_stream_ID

8位字段，定义了stream_ID的值，如表13-40所示，该字段位于包含比特流的PES包的PES包

659

头中。

表13-45　常用的stream_type码字。www.atsc.org网站的Code Point Registry提供stream_type编码的完整列表

流　类　型	编　码	流　类　型	编　码
保留	0000 0000	MPEG-4.2视觉	0001 0000
MPEG-1.2视频	0000 0001	MPEG-4.3音频	0001 0001
MPEG-2.2视频	0000 0010	由PES包传送的MPEG-4 SL打包流或FlexMux流	0001 0010
MPEG-1.3音频	0000 0011	由MPEG-4分段传送的MPEG-4 SL打包流或FlexMux流	0001 0011
MPEG-2.3音频	0000 0100	MPEG-2.6同步下载协议	0001 0100
MPEG-2私有分段	0000 0101	由PES包传送的元数据	0001 0101
包含私有数据的MPEG-2 PES包	0000 0110	由元数据分段传送的元数据	0001 0110
ISO/IEC 13552 MHEG	0000 0111	由MPEG-2.6数据轮播传送的元数据	0001 0111
MPEG-2 DSM CC	0000 1000	由MPEG-2.6对象轮播传送的元数据	0001 1000
MPEG-1、MPEG-2辅助流	0000 1001	由MPEG-2.6同步下载协议传送的元数据	0001 1001
MPEG-2.6类型A	0000 1010	MPEG-2.11 IPMP流	0001 1010
MPEG-2.6类型B	0000 1011	MPEG-4.10视频	0001 1011
MPEG-2.6类型C	0000 1100	保留	0001 1100~0111 1110
MPEG-2.6类型D	0000 1101	IPMP流	0111 1111
MPEG-2辅助流	0000 1110	用户私有	1000 0000~1111 1111
MPEG-2.7音频	0000 1111		
某些常用的用户私有细节			
DigiCipher II视频	1000 0000	ATSC数据服务表、网络资源表	1001 0101
Dolby Digital音频	1000 0001	SCTE IP数据	1010 0000
SCTE标准字幕	1000 0010	ATSC同步数据流或SCTE同步数据	1100 0010
SCTE同步数据	1000 0011	SCTE异步数据	1100 0011
ATSC节目标识符	1000 0101		
Dolby Digital Plus音频	1000 0111		

660

14. elementary_stream_info_length

16位二进制数，定义了紧跟在该字段之后的描述符总长（单位是字节）。

15. descriptor_loop

descriptor_loop中可能出现不同类型的描述符。

16. CRC_32

32位CRC，用于整个节目流映射。

13.7.4　节目流目录

节目流目录（Program Stream Directory，PSD）提供关于节目流中的比特流以及比特流相互之

间关系的描述。如果stream_ID = program stream map，PSD以PES包数据的形式出现。

13.8 传输流

传输流设计用于易出错的环境中，如远距离传输或嘈杂环境中，它被ARIB、ATSC、DVB、数字有线电视以及OpenCable™标准所采用。

传输流将一个或多个节目以及一个或多个独立时间基组合成一个单一流。每个传输流中的节目都能拥有属于自己的时间基。同一传输流中不同的节目，其时间基可能各不相同。

传输流由一个或多个188B的包组成。每个包的数据来自于PES包、PSI（节目相关信息）分段、填充字节或私有数据。除了MPEG-2数据之外，MPEG-4.2、MPEG-4.10（H.264）、SMPTE 421M（VC-1）以及其他类型数据都可采用MPEG-2传输流传送。

在每个包的开始是一个包标识符（PID），PID能够使解码器确定如何处理该数据包。如果MPEG数据采用“单载波多通道”传送，解码器通过PID确定哪些包是当前观察或记录通道的一部分，从而处理这些包，而丢弃其他包。系统信息（SI）（如节目指南、通道频率等）同样被分配唯一的PID值。

图13-23给出了传输流的一般格式。注意，起始码（0x000001xx）必须是字节对齐的，可以通过在起始码之前插入0~7个0位来实现。

13.8.1 包层

每个包由一个包头跟一个可选的适配域和一个或多个数据包组成。

1. sync_byte

8位字符串，值为0100 0111。

2. transport_error_indicator

标志位。值为1表示包中出现至少一个不可纠正的比特错误。

3. payload_unit_start_indicator

该标志位的含义取决于有效载荷。

对于PES包数据，该字段值为1表示包的数据块以PES包的字节开始。值为0表示包的数据块中没有以PES包开始的。 661

对于PSI数据，该字段值为1表示包的数据块包含PSI分段的第一个字节。

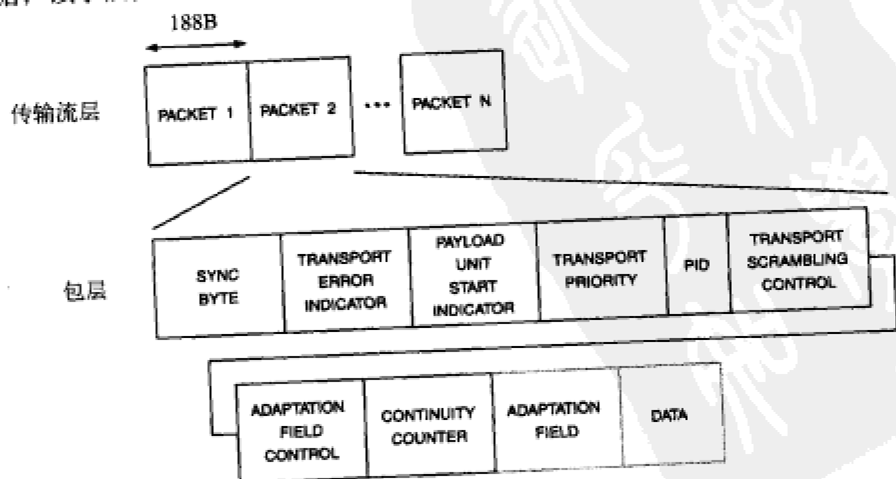


图13-23 MPEG-2传输流结构。没有给出标记和保留位。某些应用增加一个4字节TP_extra_header 前缀（产生192B），由到达时间戳和复制允许信息组成

4. transport_priority

标志位。值为1表示与其他具有相同PID的包相比，这个包拥有更高的优先级。对于DVB标准，该字段是可选的，而且即使出现也可被解码器忽略。

5. PID

13位码字，指示数据块中的数据类型，如表13-46所示。www.atsc.org网站的Code Point Registry中提供PID编码的完全列表。

6. transport_scrambling_control

2位编码，指示有效载荷的加扰模式。00＝不加扰，01＝不加扰（私有用途），10＝偶数键加扰，11＝奇数键加扰。当该字段取00之外的值时，要求在基本码流中出现CA解码器。

表13-46　常用PID码字

描　　述	编　　码
节目关联表	0 0000 0000 0000
条件接收表	0 0000 0000 0001
传输流描述表	0 0000 0000 0010
IPMP控制信息表	0 0000 0000 0011
MPEG-2保留	0 0000 0000 0100~0 0000 0000 1111
由DVB使用	0 0000 0001 0000~0 0000 0001 1111
由ARIB使用	0 0000 0010 0000~0 0000 0010 1111
由ATSC、CEA和SCTE使用	1 1111 1111 0111~1 1111 1111 1110
MPEG-2空包	1 1111 1111 1111

662

7. adaptation_field_control

2位编码，指示包头之后出现适配域还是数据块，如表13-47所示。

表13-47　MPEG-2 adaptation_field_control码字

描　　述	编　　码
保留	00
仅数据	01
仅适配域	10
适配域跟数据	11

8. continuity_counter

4位二进制数，随PID相同的包数目增加而增加。达到最大值后，它归零重新计数。当包中不存在数据块时该字段值不会增大。

9. adaptation_field

参见13.8.2节。

10. data_byte

[*n*]数据字节，是来自于PES包、PSI分段、填充字节或私有数据的毗邻数据字节。[*n*]等于184减去适配域（如果出现）中的数据字节个数。当adaptation_field_control＝01或11时该字段才出现。

13.8.2　适配域

这个域包含42位节目时钟基准（PCR）、随机访问指示符、接合倒计时以及其他MPEG和私有信息。在解码器中，PCR用来重新生成27MHz时间基时钟，该时钟与编码器中使用的相同。这个时钟也是显示时间戳（PTS）的基础。在流中PCR一般每0.1秒出现一次。当adaptation_field_control＝01或11时该域才出现。

1. adaptation_field_length

8位二进制数，定义了该字段后紧随出现的字节个数。值0000 0000用来在传输流包中插入单个的填充字节。当adaptation_field_control＝11时，该字段的值范围为0~182_D。当adaptation_

field_control = 10时，该字段的值为183_D。

注意：如果adaptation_field_length = 0000 0000，以下字段不会出现。

2. discontinuity_indicator

1位标志位。值为1表示当前传输流包中存在不连续状态。

3. random_access_indicator

1位标志位，用来指示当前传输流包以及后续可能出现的具有相同PID的传输流包是否包含帮助随机访问的某些信息。

4. elementary_stream_priority_indicator

1位标志位。在具有相同PID的包中，该字段用来指示传输流包中携带基本码流数据的优先级。值为1表示该有效载荷的优先级高于其他传输流包的有效载荷。值为0表示该有效载荷与其他使这个标志位值不为1的包具有相同的优先级。

663

5. PCR_flag

1位标志位。值为1表示适配域包含PCR字段。值为0表示适配域不包含PCR字段。

6. OPCR_flag

可选的1位标志位。值为1表示适配域包含OPCR字段。值为0表示适配域不包含任何OPCR字段。

7. splicing_point_flag

1位标志位。值为1表示splice_countdown字段在关联的适配域中出现，指定了接合点的出现。

8. transport_private_data_flag

1位标志位。值为1表示适配域包含一个或多个私有字节。

9. adaptation_field_extension_flag

1位标志位。值为1表示存在适配域的扩展。

10. program_clock_reference_base

可选42位program_clock_reference字段的33位LSB。当PCR_flag = 1时该字段才出现。

11. reserved_bits

6位可选保留位，值为111111。当PCR_flag = 1时该字段才出现。

12. program_clock_reference_extension

9位可选字段，与program_clock_reference_base字段一起组成42位节目时钟基准（PCR）。PCR指示字节预期到达的时间，该字节包含作为解码器输入的program_clock_reference_base字段的最后一位。当PCR_flag = 1时该字段才出现。

13. original_program_clock_reference_base

可选42位original_program_clock_reference字段的33位LSB。当OPCR_flag = 1时该字段才出现。

14. reserved_bits

6位可选保留位，值为111111。当OPCR_flag = 1时该字段才出现。

15. original_program_clock_reference_extension

9位可选字段，与original_program_clock_reference_base字段一起组成42位原始节目时钟基准（OPCR）。OPCR仅在包含PCR的传输流包中出现。OPCR支持单一节目传输流从另一传输流中重建。当OPCR_flag = 1时该字段才出现。

16. splice_countdown

2的反码，长度为8位，代表一个可正可负的值。正值定义了在接合点之前与关联传输流包具

664 有相同PID值的传输流包个数。负值指示在接合点之后的关联传输流包是第n个包。当splicing_point_flag＝1时这个字段才出现。

17. transport_private_data_length

8位可选的二进制数[n]，定义了这个字段后紧接着出现的字节数。当transport_private_data_flag＝1时该字段才出现。

18. private_data_byte

可选的[n]数据字节，MPEG-2没有定义该字段。当transport_private_data_flag＝1时该字段才出现。

19. adaptation_field_extension_length

可选的8位二进制数，指示这个字段后紧接着出现的扩展适配域数据的字节数，如果出现，还包括保留字节。当adaptation_field_extension_flag＝1时该字段才出现。

20. ltw_flag

1位可选标志位（合法的时间窗口标志），值为1指示ltw_offset字段的出现。当adaptation_field_extension_flag＝1时该字段才出现。

21. piecewise_rate_flag

1位可选标志位，值为1指示piece_rate字段的出现。当adaptation_field_extension_flag＝1时该字段才出现。

22. seamless_splice_flag

1位可选的标志位，值为1指示splice_type和DTS_next_AU字段的出现。当adaptation_field_extension_flag＝1时该字段才出现。

23. reserved_bits

5位可选保留位，值为11111。当adaptation_field_extension_flag＝1时该字段才出现。

24. ltw_valid_flag

1位可选标志位，指示ltw_offset的值是否有效。当adaptation_field_extension_flag＝1以及ltw_flag＝1时该字段才出现。

25. ltw_offset

15位可选的二进制数，定义了单位为（$300/f_s$）秒的合法时间窗口偏移，这里f_s是该PID所属节目的系统时钟频率。当adaptation_field_extension_flag＝1以及ltw_flag＝1时该字段才出现。

26. reserved_bits

2位可选保留位，值为11。当adaptation_field_extension_flag＝1以及piecewise_rate_flag＝1时该字段才出现。

27. piecewise_rate

可选的22位二进制数，定义了一个假定的比特率，该比特率用来界定具有相同PID值的传输流包中合法时间窗口的结束时间，这些传输流包在当前包之后出现且不包含legal_time_window_offset字段。当adaptation_field_extension_flag＝1以及piecewise_rate_flag＝1时该字段才出现。

665

28. splice_type

4位可选的二进制数，在它出现的所有后续具有相同PID值的传输流包中，该字段具有相同的值，直到包中的splice_countdown字段值归零为止（包括当前包）。如果带入该PID的基本流是音频流，该字段的值将为0000。如果带入该PID的基本流是视频流，该字段指示将引起用于接合的基本流注意的状态。当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时该字段才出现。

29. DTS_next_AU[32...30]

可选的33位字段，指示接合点后第一个访问单元的解码时间。当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时该字段才出现。

30. marker_bit

可选的1位标记位。值恒为1。当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时该字段才出现。

31. DTS_next_AU[29...15]

可选字段，当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时才出现。

32. marker_bit

可选的1位标记位。值恒为1。当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时该字段才出现。

33. DTS_next_AU[14...0]

可选字段，当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时才出现。

34. marker_bit

可选的1位标记位。值恒为1。当adaptation_field_extension_flag＝1以及seamless_splice_flag＝1时该字段才出现。

35. reserved_bits

可选的[*n*]数据字节，值为1111 1111。当ltw_flag、piecewise_rate_flag以及seamless_splice_flag＝0，而adaptation_field_extension_flag＝1时，该字段才出现。

36. stuffing_byte

可选的[*n*]数据字节，值为1111 1111。当OPCR_flag、adaptation_field_extension_flag、PCR_flag、transport_private_data_flag以及splicing_point_flag＝0时该字段才出现。

13.8.3 节目相关信息

节目相关信息（PSI）是一类附加信息，使解码器能够更加高效地在单一传输流中找出所需的内容，并集成为用户友好的电子节目指南（EPG）。

节目由一个或多个的打包音频、视频、数据等基本流组成，每个流都分配一个13位的包标识号（PID）。此外，携带相同PES的传输流包分配相同但唯一的PID。

只有当获得正确的PID后MPEG-2解码器才能处理正确的包。这是PSI的功能。传输流包中携带的PSI具有唯一的PID，因而MPEG-2解码器能够方便地找到它。

666

PSI包括6个表，如表13-48所示。节目关联表（PAT）、条件接收表（CAT）、传输流描述表（TSDT）、IPMP控制信息表（ICIP）以及空包是仅有的固定的PID。MPEG-2解码器通过访问适当的表来确定剩余的PID。

如果首次接收传输流，MEPG-2解码器将查找PAT、CAT、TSDT和ICIT。如图13-24所示，从PIA中读出网络信息表（NIT）的PID和每个PMT。再从PMT中读出每个基本流的PID。如果节目是加密的，还要求访问CAT。

表13-48 节目相关信息表

流 类 型	缩 写	PID	描 述
节目关联表	PAT	0x0000	关联节目编号与节目映射表PID
条件接收表	CAT	0x0001	关联一个或多个（私有）EMM流，每一个都有一个唯一的PID值
传输流描述表	TSDT	0x0002	关联一个或多个描述符到一个完整的传输流

（续）

流 类 型	缩 写	PID	描 述
IPMP控制信息表	ICIT	0x0003	包含MPEG-2.11中定义的IPMP工具列表、权利箱、工具箱
节目映射表	PMT	由PAT分配	为一个或多个节目的组件定义PID值
网络信息表	NIT	由PAT分配*	物理网络参数,如FDM频率、转发器数量等

* 对于ARIB和许多DVB系统1PID＝0x0010。

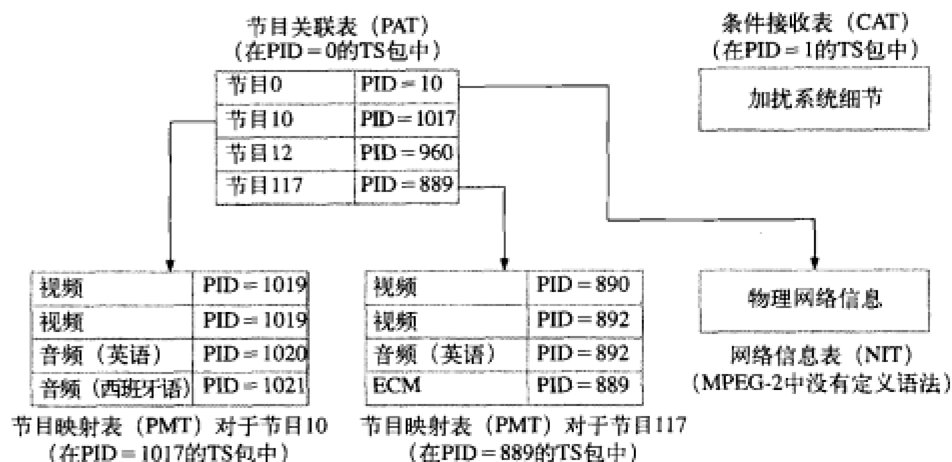


图13-24 MPEG-2 PAT和PMT示例

13.8.4 节目关联表

每个传输流包含一个或多个PID值为0x0000的传输流包。所有这些包一起组成了完整的节目关联表（PAT）。

如图13-24所示，PAT提供传输流中所有节目的完全列表。每个节目都包括传输流包的PID值，这些包中包含它对应的节目映射表（PMT）。

667
~
668

由于没有PAT就不可能分离信号，新节目的锁定速度取决于PAT传送的频率。MPEG-2在PAT和任意在PAT中提到的PMT之间定义了一个最大为0.5秒的间隔值。

采用下列语法，PAT在插入传输流包前可以分割成一个或多个分段（section）。

1. table_ID

8位码字，标识了内容的类型，值为0x00，如表13-49所示。www.atsc.org网站的Code Point Registry中提供table_ID编码的完全列表。

表13-49 常用的Table_ID码字

描 述	编 码
PAT分段	0x00
CAT分段	0x01
PMT分段	0x02
TSDT分段	0x03
MPEG-4场景描述分段	0x04
MPEG-4对象描述符分段	0x05
元数据分段	0x06
ICIT分段	0x07

（续）

描述	编码
MPEG-2保留	0x08~0x38
MPEG-2.6 DSM-CC可寻址分段	0x39
包含多协议数据的MPEG-2.6 DSM-CC分段	0x3A
包含U-N消息的MPEG-2.6 DSM-CC分段	0x3B
包含下载数据消息的MPEG-2.6 DSM-CC分段	0x3C
包含流描述符的MPEG-2.6 DSM-CC分段	0x3D
包含私有数据的MPEG-2.6 DSM-CC分段	0x3E
MPEG-2.6 DSM-CC可寻址分段	0x3F
由DVB使用	0x40~0x7F
由ARIB、ATSC CA和DVB CA所使用	0x80~0x8F
由ATSC和SCTE使用	0xC0~0xFE
禁用	0xFF

2. section_syntax_indicator

1位标志位，值恒为1。

3. reserved_bits

2位保留位，值为01。

4. section_length

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过1021_D或0x3FD。

5. transport_stream_ID

16位二进制数，作为标签标识网络中来自于其他多路通道的传输流。

6. reserved_bits

2位保留位，值为11。

7. version_number

5位二进制数，它是整个PAT的版本号。任何时候只要PAT的定义发生变化，版本号就加1（模数32）。

当current_next_indicator值为1时，version_number将是当前可用PAT的版本号。当current_next_indicator值为0时，version_number将是下一个可用PAT的版本号。

8. current_next_indicator

1位标志位，值为1表示PAT传送当前可用。值为0表示PAT传送还不可用，但下一个PAT传送将变为有效。

9. section_number

8位二进制数，定义了分段的编号。PAT中第一分段的section_number值必须为0x00。每当出现附加分段，PAT中该字段的值就加1。

10. last_section_number

8位二进制数，定义了完整PAT最后一个分段的编号（也就是section_number值最高的分段）。

注意：对于每个节目，下列4个字节会重复出现。

11. program_number

16位二进制数，定义了program_map_PID分配的节目。

12. reserved_bits

3位保留位，值为111。

13. network_PID

13位二进制数，定义了包含网络信息表的传输流包的PID。它的值可能为0x0010~0x1FFE。仅当program_number＝0x0000时该字节才出现。

14. program_map_PID

13位二进制数，定义了包含PMT分段的传输流包的PID，对于program_number定义的节目，这个PMT分段是可用的。它的值可能为0x0010~0x1FFE。仅当program_number ≠ 0x0000时该字节才出现。

15. CRC_32

669

32位CRC值。

13.8.5　节目映射表

如图13-24所示，节目映射表（PMT）提供了节目编号和节目元素（视频、音频等）之间的映射关系。这项工作通过标识音频、视频和其他属于给定节目的流的PID值完成。

需要注意的是，对于ATSC、DVB和OpenCable™标准，某些PMT要求特定的PID。因此，MPEG-2和DVB/ATSC/OpenCable™比特流不能完全交互。

采用下列语法，PMT仅有一个由program_number标识的分段。

1. table_ID

8位码字，标识了内容的类型，值为0x02，如表13-49所示。www.atsc.org网站的Code Point Registry中提供table_ID编码的完全列表。

2. section_syntax_indicator

1位标志位，值恒为1。

3. reserved_bits

3位保留位，值为011。

4. section_length

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过1021_D或0x3FD。

5. program_number

16位二进制数，定义了program_map_PID适用的节目。一个节目映射分段传送一个节目定义。节目定义绝不能长于1016_D（0x3F8）。

6. reserved_bits

2位保留位，值为11。

7. version_number

5位二进制数，它是节目映射分段的版本号。当分段出现时携带的信息发生变化，版本号加1（模数32）。

当current_next_indicator值为1时，version_number是当前可用节目映射分段。当current_next_indicator值为0时，version_number将是下一个可用节目映射分段。

8. current_next_indicator

1位标志位，值为1表示节目映射分段传送当前可用。值为0表示节目映射分段传送还不可用，但下一个节目映射分段传送变为有效。

9. section_number

8位二进制数，值为0x00。

10. last_section_number

8位二进制数，值为0x00。

11. reserved_bits

3位保留位，值为111。

670

12. PCR_PID

13位二进制数，指示包含PCR字段传输流包的PID，对于program_number定义的节目，这一PCR字段是有效的。这一字段的值可能为0x0010~0x1FFE。如果对于私有流，没有PCR与节目定义关联，那么这一字段的值为0x1FFF。

13. reserved_bits

4位保留位，值为1111。

14. program_info_length

12位二进制数，定义了紧接该字段之后出现的描述符总长（单位为字节）。

15. descriptor_loop

[n]描述符，可能在descriptor_loop中出现。

注意：对于任意出现的流类型，下列6个字段会重复出现。

16. stream_type

8位码字，定义了包中节目元素的类型，该包的PID由elementary_PID定义。stream_type的值在表13-45中定义。

17. reserved_bits

3位保留位，值为111。

18. elementary_PID

13位二进制数，定义了携带相关节目元素的传输流包的PID。

19. reserved_bits

4位保留位，值为1111。

20. ES_info_length

12位二进制数，定义了紧接该字段后出现的描述符总长（单位为字节）。

21. descriptor_loop

[n]描述符，可能在descriptor_loop中出现。

22. CRC_32

32位CRC值。

13.8.6 传输流描述表

可选的传输流描述表（TSDT）用来包含应用于整个传输流的描述符。

采用下列语法，在插入传输流包之前，TSDT可以分割成一个或多个分段。

1. table_ID

8位码字，标识了内容的类型，值为0x03，如表13-49所示。www.atsc.org网站的Code Point Registry中提供table_ID编码的完全列表。

2. section_syntax_indicator

1位标志位，值恒为1。

3. reserved_bit

1位保留位，值为0。

4. reserved_bits

671

2位保留位，值为11。

5. section_length

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过1021_D或0x3FD。

6. reserved_bits

18位保留位，值为11 1111 1111 1111 1111。

7. version_number

5位二进制数，它是节目映射分段的版本号。当分段出现时携带的信息就要发生变化，版本号加1（模数32）。

当current_next_indicator值为1时，version_number是当前可用节目映射分段。当current_next_indicator值为0时，version_number将是下一个可用节目映射分段。

8. current_next_indicator

1位标志位，值为1表示节目映射分段传送当前可用。值为0表示节目映射分段传送还不可用，但下一个节目映射分段传送变为有效。

9. section_number

8位二进制数，定义了分段的编号。第一个分段的section_number值必须为0x00。每当出现附加的分段时，该字段的值就加1。

10. last_section_number

8位二进制数，定义了最后一个分段的编号（也就是section_number值最高的分段）。

11. descriptor_loop

[*n*]描述符，可能在descriptor_loop中出现。

12. CRC_32

32位CRC值。

13.8.7 条件接收表

条件接收表（CAT）提供一个或多个条件接收（CA）系统、它们的权利管理消息（EMM）以及与之相关的特定参数之间的关联。CAT还包含权利控制消息（ECM）和权利管理消息（EMM）的PID。

采用下列语法，在插入传输流包之前，CAT可以分割成一个或多个分段。

1. table_ID

8位码字，标识了内容的类型，值为0x01，如表13-49所示。www.atsc.org网站的Code Point Registry中提供table_ID编码的完全列表。

2. section_syntax_indicator

1位标志位，值恒为1。

3. reserved_bits

3位保留位，值为011。

4. section_length

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字

672

段的值不能超过1021_D或0x3FD。

5. reserved_bits

18位保留位，值为11 1111 1111 1111 1111。

6. version_number

5位二进制数，它是整个CAT的版本号。当CAT变化时携带的信息发生变化，版本号加1（模数32）。

当current_next_indicator值为1时，version_number将是当前可用CAT。当current_next_indicator值为0时，version_number将是下一个可用CAT。

7.current_next_indicator

当1位标志位为1时，发送的CAT当前可用。当1位标志位为0时，发送的CAT尚不可用，下一个CAT才有效。

8. section_number

8位二进制数，定义了分段的编号。CAT中第一个分段的section_number值必须为0x00。每当CAT中出现附加分段，该字段的值就加1。

9. last_section_number

8位二进制数，定义了整个CAT最后一个分段的编号（也就是section_number值最高的分段）。

10. descriptor_loop

[*n*]描述符，可能在descriptor_loop中出现。

11. CRC_32

32位CRC值。

13.8.8 网络信息表

PAT中的首次登录是为网络数据保留的，且包含网络信息表（NIT）的PID。NIT包括其他可能可用传输流的信息，例如通过调谐至不同的RF频道或卫星。每个传输流可能包括定义射频、卫星轨道位置等信息的描述符。在MPEG-2中，对于这个目的只有NIT是强制要求的。在DVB中采用称为DVB-SI的附加元数据，而且NIT被视作DVB-SI的一部分。因此，术语PSI/SI是通用术语。

13.8.9 IPMP控制信息表

IPMP（知识产权管理和保护）控制信息表包含IPMP[也被称为数字权利管理(DRM）]相关信息，其中包括工具列表、权利箱和工具箱。

*工具列表*表示并且启用处理内容所需的IPMP工具的选择。*工具箱*使内容流中能够携带二进制工具。*权利箱*可能包含权利描述，这些描述说明了与IPMP保护内容相关的适用规则。

IPMP流携带的所有类型IPMP数据（包括健、ECM、EMM）都会传送给这些工具。

673

13.9 知识产权管理和保护

知识产权管理和保护（IPMP）也被称作数字权利管理（DRM），它为知识产权管理的实现提供一类接口和工具，而不是一个完整的系统。

已提供的管理与保护的层次和类型取决于内容和业务模型的值。因此，IPMP系统的完整设计留待应用开发者来完成。

这一框架保证了开放和私有解决方案的使用，但同时保证互操作性，支持多种类型保护（如解密、水印、权利管理等）的使用，同时支持采用定义过的设备间消息（反映在家庭网络上的内

容分发问题）实现设备间的内容传送。

对于受保护的内容，IPMP工具要求在显示开始前与解码器进行通信。工具配置和初始化信息通过IPMP描述符或IPMP基本流传送。所需工具可以通过内嵌、下载或其他方法获得。

在IPMP描述符中的控制点和顺序序列信息允许在系统的不同位置使用不同工具。IPMP描述符或IPMP基本流携带的IPMP数据包括权利箱、键箱和工具初始化数据。

13.10　在MPEG-2传输流上的MPEG-4.2视频

MPEG-2传输或节目流取代MPEG-2视频来传输MPEG-4.2视频。这使得现有的基础设施和设备能够方便地适应MPEG-4.2视频编解码。

对于MPEG-4.2视频，PES包的stream_id＝1110 xxxx。PMT或PSM中的stream_type＝0x10。MPEG-4.2流的传送也必须采用MPEG-4*视频描述符*发出信号告知。

MPEG-4.3音频、MPEG-4 SL打包流以及MPEG-4 FlexMux流也能够通过MPEG-2传输或节目流传送。

13.11　在MPEG-2传输流上的MPEG-4.10（H.264）

MPEG-2传输或节目流取代MPEG-2视频在PES包中传送MPEG-4.10（H.264）视频。这使得现有的基础设施和设备能够方便地适应H.264视频编解码。

对于MPEG-4.10（H.264）视频，PES包的stream_id＝1110 xxxx。PMT或PSM中的stream_type＝0x1B。H.264流的传送也必须采用MPEG-2 AVC*视频描述符*或MPEG-2 AVC*时序及*HRD*描述符*发出信号告知。

674

13.12　在MPEG-2传输流上的SMPTE 421M（VC-1）视频

MPEG-2传输和节目流取代MPEG-2视频在PES包中传送SMPTE 421M（VC-1）视频。这使得现有的基础设施和设备能够方便地适应SMPTE 421M（VC-1）视频编解码。

对于SMPTE 421M视频，PES包stream_ID＝0xFD，stream_ID_extension＝101 0101。在PMT或者PSM中，stream_type＝0xEA。SMPTE 421M流的传送必须采用一个或多个MPEG-2*注册描述符*（Registration Descriptor）来发出信号告知，每一个描述符可能在additional_identification_info字段中包含一个可选的SMPTE 421M子描述符。

SMPTE 421M*档次和级别子描述符*（Profile and Level Subdescriptor）可能被用来指定SMPTE 421M流的档次和级别。SMPTE 421M*对齐子描述符*（Alignment Subdescriptor）可能被用来定义在编码的字节序列和一个PES包之间存在哪一种对齐类型。SMPTE 421M*缓冲大小子描述符*（Buffer Size Subdescriptor）可能被用来定义在解码器中用于解码SMPTE 421M流所需的最小基本流缓冲区大小。

对于Simple或者Main档次流，VC-1_SPMP_PESpacket_PayloadFormatHeader()必须出现在每一个存取单元的起始处。

13.13　MPEG-2 PMT/PSM描述符

这些MPEG-2描述符被用来标识出现在MPEG-2传输或者节目流中的常用私有（非MPEG-2）信息。

描述符通常包含在MPEG-2 PMT或者PSM中的descriptor_loop中。描述符的一般格式是：

descriptor_tag（8位）

descriptor_length（8位）

data

描述符标签值为0和1，而44~63保留。值19~26为MPEG-2.6数据保留。除非有其他指示，描述符可以同时出现在传输和节目流中。

13.13.1 MPEG-2 AAC音频描述符

对于在PES包中传送的独立MPEG-2.7音频流，MPEG-2 AAC描述符为标识这个音频基本流的编码参数提供了基本的信息。

1. descriptor_tag

长度为8位的字段，值为0010 1011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0011。

3. MPEG-2_AAC_profile

8位字段，标识每个MPEG-2.7规范的AAC档次。

675

4. MPEG-2_AAC_channel_configuration

8位字段，标识音频信道的数目和配置。

5. MPEG-2_AAC_additional_configuration

8位字段，标识是否有带宽扩展数据被嵌入到音频流中。

13.13.2 音频流描述符

音频流描述符提供了标识一个音频基本流编码版本的基本信息。

1. descriptor_tag

8位字段，值为0000 0011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

3. free_format_flag

1位标志位，值为1指示音频流中的bitrate_index字段是0000。

4. ID

这一位设置与音频流中的ID字段相同的值。

5. layer

2位的二进制数，设置与任何音频流中的最高层相同或者更高的值。

6. variable_rate_audio_indicator

值为0表示音频流的比特率在音频帧之间不发生变化。

7. reserved_bits

3位保留位，值恒为111。

13.13.3 AVC时序和HRD描述符

对于在PES包中传送的独立MPEG-4.10（H.264）视频流，这个MPEG-2描述符描述了视频流时间信息和假定的参考解码器（hypothetical reference decoder，HRD）信息。当H.264视频流没有传递H.264视频可用信息（video usability information，VUI）参数时，这个描述符必须出现在PMT中。

1. descriptor_tag

8位字段，值为0010 1010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. HRD_management_valid_flag

1位标志位，值为1指示缓冲周期SEI（参看H.264）需要包含在H.264视频流中。值为0指示必须采用MPEG-2漏斗方法。

4. reserved_bits

6位保留位，值恒为11 1111。

5. picture_and_timing_info_present

值为1表示该描述符包含90kHz_flag和用于映射到系统时钟的参数。

6. 90kHz_flag

值为1表示H.264时间基是90kHz。0表示N和K字段将出现。仅当picture_and_timing_info_present＝1时这个字段才出现。

676

7. reserved_bits

7位保留位，值恒为111 1111。仅当picture_and_timing_info_present＝1时这个字段才出现。

8. N，K

长度为32位的字段，描述了H.264的time_scale和system_clock_reference之间的关系。仅当picture_and_timing_info_present＝1并且90kHz_flag＝0时这些字段才出现。

9. num_units_in_tick

长度为32位的字段，其定义请参看第14章。仅当picture_and_timing_info_present＝1时这个字段才出现。

10. fixed_frame_rate_flag

1位标志位，值为1表示H.264编码视频帧率为固定值。值为0表示在此描述符中没有关于帧率的任何信息。

11. temporal_proc_flag

1位标志位，该位值为1且fixed_frame_rate_flag＝1表示H.264视频流必须传送图像顺序计数（picture order count，POC）信息。该位值为0表示此描述符中没有任何关于H.264视频流的POC信息与时间之间关系的信息。

12. picture_to_display_conversion_flag

1位标志位，值为1表示H.264视频流包含关于显示编码图像的信息。当这一位值为0时，H.264视频流的pic_struct_present_flag必须设置为0。

13. reserved_bits

5位保留位，值恒为1 1111。

13.13.4　AVC视频描述符

对于在PES包中传送的独立MPEG-4.10（H.264）视频流，这个MPEG-2描述符描述了视频流的编码参数。当这个描述符没有出现在PMT中时，流应该不包含H.264静态图像或者H.264 24小时图片。

1. descriptor_tag

8位字段，值为0010 1000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。

3. profile_IDC

8位字段，标识了遵循H.264规范的H.264视频流的档次（profile）。

4. constraint_set0_flag

1位标志位，其定义在H.264规范中给出。

5. constraint_set1_flag

1位标志位，其定义在H.264规范中给出。

6. constraint_set2_flag

1位标志位，其定义在H.264规范中给出。

7. AVC_compatible_flags

5位标志位，与H.264规范中的序列参数集中的reserved_zero_5bits有相同的值。 677

8. level_IDC

8位字段，标识了遵循MPEG-4.10（H.264）规范的MPEG-4.10（H.264）视频流的级别（level）。

9. AVC_still_present

这一位为1表示H.264视频流包含H.264静态图像。值为0表示H.264视频流不应该包含H.264静态图像。

10. AVC_24_hour_picture_flag

1位标志位，值为1表示H.264视频流包含24小时图片，这些图片是显示时间超过24小时的存取单元。值为0表示H.264视频流不应该包含24小时图片。

11. reserved_bits

6位保留位，值恒为11 1111。

13.13.5 CA描述符

这个MPEG-2描述符表示包含ECM、EMM或者SRM信息的传输流包的PID。如果出现在CAT中，那么存在一个系统范围内的条件接收管理系统。如果出现在PMT中，CA_PID指向包含与节目相关的ECM包。

1. descriptor_tag

8位字段，值为0000 1001。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. CA_system_ID

16位的二进制数，定义了条件接收使用的类型。这个字段的编码是私下定义的。值为0x4ADD标明它是一个ATSC SRM参考描述符（Reference Descriptor）。

4. reserved_bits

3位保留位，值恒为111。

5. CA_PID或SRM_PID

13位的二进制数，定义了传输流包的PID，这些传输流包或者包含ECM和EMM信息，或者包含关于由CA_system_ID定义的条件接收系统的SRM信息。

对于传输流，该字段值为0x0003表示在传输流中有组件使用IPMP。对于节目流，stream_ID_

extension值为0x00表示在节目流中有组件使用IPMP。

6. private_data_byte

这些可选的[n]字节私有数据由CA所属者定义。

13.13.6　字幕服务描述符

对于OpenCable™和一般的MPEG-2解码器，对于每一个有隐藏字幕（简称CC）的节目，这个CEA-708描述符必须出现。对于ATSC和OpenCable™，这个描述符同样也必须出现在事件信息表（event information table，EIT）的descriptor_loop中。

MPEG-2最多支持16个独立的服务描述符，每一个服务描述符长度为6字节。

1. descriptor_tag

678

8位字段，值为1000 0110。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. reserved_bits

3位保留位，值恒为111。

4. number_of_services

5位字段，取值为1~16，表示字幕服务出现的数目。

注意：[number_of_services]定义了下面9个字段重复的次数。

5. language

3位编码，指定与每个ISO 639.2/B的字幕服务相关的语言。

6. digital_CC

1位标志位，值为1表示CEA-708字幕服务出现。值为0表示CEA-608字幕出现。

7. reserved_bit

保留位，值恒为1。

8. reserved_bits

5位可选保留位，值恒为1 1111，仅当caption_type＝0时这些保留位才出现。

9. line21_field

1位可选标志位，值为1表示场2的CEA-608字幕出现。值为0表示场1的CEA-608字幕出现。仅当caption_type＝0时这一位才出现。

10. caption_service_number

可选的6位字段，取值范围为1~63，表示字幕流的服务数。仅当caption_type＝1时这些位才出现。

11. easy_reader

1位标志位，值为1表示字幕服务包含了为新读者（beginning reader）提供的文本格式。值为0表示字幕服务不是为此制作的。

12. wide_aspect_ratio

1位标志位，值为1表示字幕服务是为16：9显示器而格式化的。值为0表示字幕服务是为4：3显示器而格式化的，且可以可选地显示在16：9显示器的中心。

13. reserved_bits

14位保留位，值恒为11 1111 1111 1111。

13.13.7 版权描述符

这个MPEG-2描述符提供了一种为音视频作品标识身份的方法。对于DVB标准，这个描述符是可选的，如果存在，解码器可以将其忽略。

1. descriptor_tag

8位字段，值为0000 1101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

679

3. copyright_ID

32位值，从版权登记机构获得。

4. additional_copyright_info

可选的[n]字节数据，由版权所有者自行定义，永不变更。

13.13.8 数据流对齐描述符

这个MPEG-2描述符描述了相对于PES包载荷起始点的视频流语法的对齐准则。ATSC需要这个描述符出现在PMT分段的节目元素循环中，用来描述视频基本流。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 0110。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

3. alignment_type

8位的码字，定义了音频或者视频对齐类型，如表13-50所示。对于ATSC标准，这个值必须是0000 0010。

表13-50 alignment_type码字

视频流对齐类型	音频流对齐类型	码
保留	保留	0000 0000
条带、图像、GOP或者序列	同步字	0000 0001
图像、GOP或序列	保留	0000 0010
GOP或序列		0000 0011
序列		0000 0100
保留		0000 0101
		⋮
		1111 1111

13.13.9 DTCP描述符

这个描述符用来控制HDCP保护和DTCP保护的数字输出，如IEEE 1394、USB、HDMI和IP网络。

1. descriptor_tag

8位字段，值为1000 1000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

680

3. CA_system_ID

16位的二进制数，定义了条件接收使用的类型。值为0x0FFF（DTLA）。

4. reserved_bits

5位保留位，值恒为1 1111。

5. EPN

1位加密加非声明（Encryption Plus NonAssertion，EPN）标志位，值为1说明非其他复制控制的内容将不会在Internet上重传。

6. DTCP_CCI

2位码字，指定了数字复制代次管理：

00＝自由复制

01＝不再复制

10＝复制1代

11＝从不复制

7. reserved_bits

5位保留位，值恒为1 1111。

8. image_constrain_token

1位标志位，值为1表示当输出到不受保护的高清晰模拟视频输出设备上时，高清晰内容必须限定最高总共520 000像素或者更少（例如，960×540p）。

9. APS

2位的码字，详细说明模拟保护服务（Analog Protection Service，APS）。

00＝没有模拟保护服务

01＝PSP开，颜色带关

10＝PSP开，2行颜色带开

11＝PSP开，4行颜色带开

13.13.10 DTS音频描述符

包含DTS®音频的PES包可能被包括在一个MPEG-2节目或者传输流中，其方式与MPEG或Dolby® Digital音频可以被包括的方式一样。

MPEG-2并不显式地支持DTS®比特流。同样，在PSI表中，MPEG-2音频流描述符并没有充分地描述DTS®比特流的内容。

因此，包含DTS®音频数据的PES包将采用私有流1来发送。此外，还需要一个注册描述符Registration Descriptor（descriptor_tag＝0000 0101）和一个DTS音频描述符（descriptor_tag＝1001 0001或0111 0011）。

13.13.11 层次描述符

这个MPEG-2描述符提供了标识节目元素的信息，这些节目元素包含了层次编码视频、音频和私有流的组件。

1. descriptor_tag

8位字段，值为0000 0100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。

3. reserved_bits

4位保留位，值恒为1111。

4. hierarchy_type

4位码字，指示层次层与其嵌入的层之间的关系，如表13-51所示。

表13-51 hierarchy_type码字

层次类型	编码
保留	0000
空域可分级	0001
SNR可分级	0010
时域可分级	0011
数据分区	0100
扩展比特流	0101
私有比特流	0110
多视角档次	0111
保留	1000~1110
基本层	1111

681

5. reserved_bits

2位保留位，值恒为11。

6. hierarchy_layer_index

6位二进制数，表示流的一个唯一索引。

7. reserved_bits

2位保留位，值恒为11。

8. hierarchy_embedded_layer_index

6位二进制数，定义了必须在解码之前访问的流层次表索引。这个参数对于hierarchy_type＝1111是未定义的。

9. reserved_bits

2位保留位，值恒为11。

10. hierarchy_channel

6位二进制制数，表示计划的通道编号，最健壮的通道有最低的值。

13.13.12 IBP描述符

这个MPEG-2描述符提供了关于视频序列中帧类型序列的某些特征的信息。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0001 0010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0010。

3. closed_gop_flag

1位标志位，值为1表示在每一个I帧之前都有一个编码的图像组头，且在视频序列中所有图像组头中的closed_gop标志位都设置为1。

4. identical_gop_flag

1位标志位，表示I帧之间的P帧和B帧数目以及I图像之间的图像编码类型和图像类型序列数目在整个序列中都是一致的，只有两个I图像情况可能例外。

5. max_gop_length

14位的二进制数，表示视频序列中任意两个连续I图像之间的最大编码图像数。值为0被禁止使用。

682

13.13.13 IPMP描述符

这个MPEG-2描述符告知IPMP工具保护，将IPMP工具与每一个受保护的节目相关联，指示特定IPMP工具应该运行的控制点。

MPEG-2 IPMP描述符的descriptor_tag＝0010 1001。

13.13.14　ISO 639语言描述符

这个MPEG-2描述符提供了指示每一个音频基本流语言的方法。如果存在，对于每一个Dolby® Digital或Dolby® Digital Plus音频基本流，ATSC需要这个描述符出现在PMT中ES_info_length之后的descriptor_loop中。对于DVB标准，如果超过一个以上有不同语言的音频（或视频）流用于一个节目，这个描述符必须出现并且被解码。

1. descriptor_tag

8位字段，值为0000 1010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：对于每一种语言，下面2个字段都会出现。

3. ISO_639_language_code

24位字段，包含了一个3字符的语言码。

4. audio_type

8位码字，详细说明了音频类型：

0x00＝保留

0x01＝无解说词的电视背景音乐（没有语言）

0x02＝为听力障碍人士准备的字幕

0x03＝为视障人士准备的解说词

0x04~0xFF＝保留

13.13.15　最大比特率描述符

这个MPEG-2描述符提供了一个用来说明关于最大比特率出现的信息的方法。该描述符仅应用到传输流，而没有应用到节目流。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 1110。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，其值为0000 0011。

3. reserved_bits

2位保留位，值恒为11。

4. maximum_bitrate

22位的二进制数，表示最大比特率出现，单位是50字节/秒。

13.13.16　元数据描述符

683

各种MPEG-2描述符可用来包含元数据信息，这些信息描述了音视觉内容和数据本质。

MPEG-2定义了两个用于同步传输元数据的工具：PES包载荷、DSM-CC同步下载协议。

此外，MPEG-2定义了三个用于异步传输元数据的工具：在元数据分段传送、DSM-CC数据轮播、DSM-CC对象轮播。

1. 内容标签描述符

这个MPEG-2描述符分配一个标签给内容，标签可以被元数据用来指向相关的内容。该描述

符同样提供了关于采用哪一个内容时间基以及关于内容时间基与元数据时间基之间偏移的信息。

这个描述符的descriptor_tag值为0010 0100。

2. 元数据指针描述符

这个MPEG-2描述符指向一个元数据服务，并将其服务与MPEG-2流中的音视频内容关联起来。

这个描述符的descriptor_tag值为0010 0101。

3. 元数据描述符

这个MPEG-2描述符定义了由节目流或传输流传送的伴随元数据格式。该描述符也可传递信息来标识元数据服务，这些元数据服务来自一个在DSM-CC轮播上传输的元数据集合。

这个描述符的descriptor_tag值为0010 0110。

4. 元数据STD描述符

这个MPEG-2描述符定义了标准模型的参数，这些标准模型是用于元数据流的处理的。

这个描述符的descriptor_tag值为0010 0111。

13.13.17 复用缓冲区利用描述符

这个MPEG-2描述符提供了STD复用缓冲区的容量界限。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 1100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，其值为0000 0100。

3. bound_valid_flag

1位标志位，值为1表示LTW_offset_lower_bound和LTW_offset_upper_bound是有效的。

4. LTW_offset_lower_bound

15位的二进制数，单位为（27MHz/300）的时钟周期，定义了任何LTW_offset字段的最低值。

5. reserved_bit

保留位，值恒为1。

6. LTW_offset_upper_bound

15位的二进制数，单位为（27MHz/300）的时钟周期，定义了任何LTW_offset字段的最高值。 684

13.13.18 私有数据指示器描述符

这个MPEG-2描述符提供了一个传送私有信息的方法。对于DVB标准，这个MPEG-2描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 1111。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，其值为0000 0100。

3. private_data_indicator

32位的私有数据，MPEG-2规范中没有相关定义。

13.13.19 注册描述符

这个MPEG-2描述符提供了一种唯一标识私有数据格式的方法。

遵循ATSC标准的节目由这个描述符在PMT的program_info_length之后的descriptor_loop中标识。

对于每一个stream_type值在ATSC用户私有范围（oxC4~0xFF）之内的节目元素，这个描述符同样也出现在PMT的ES_info_length之后的descriptor_loop中，其目的是建立与该节目元素伴随的私有实体。

对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 0101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. format_identifier

32位格式ID标识值，从SMPTE注册机构获得。表13-52给出了一些常见的实例。

表13-52 format_identifier码字

格式	编码
ATSC	0x47413934
AVS（GB/T 20090.2-2006）	0x41565356
Dolby Digital音频	0x41432D33
DTS音频（512帧）	0x44545331
DTS音频（1024帧）	0x44545332
DTS音频（2048帧）	0x44545333
SCTE 54标准	0x53435445
SMPTE 421M（VC-1）视频	0x56432D31

4. additional_identification_info

可选的[n]字节数据，由注册所属者自己定义，且永不变更。

13.13.20 平滑缓冲区描述符

这个MPEG-2描述符传递的是关于与这个描述符相关的平滑缓冲区的大小和与之相关的缓冲区漏出速率的信息。ASTC需要PMT拥有一个与其节目相关的平滑缓冲区描述符。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

685

8位字段，值为0001 0000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0110。

3. reserved_bits

2位保留位，值恒为11。

4. sb_leak_rate

22位的二进制数，为相关的基本流或者其他数据定义了从缓冲区漏出的速率，单位为400 bit/s。

5. reserved_bits

2位保留位，值恒为11。

6. sb_size

22位的二进制数，为相关的基本流或者其他数据定义了平滑缓冲区大小（以字节为单位）。对于ATSC和OpenCable™标准，这个字段的值小于等于2048。

13.13.21 STD描述符

这个MPEG-2描述符仅应用到传输流，而没有应用到节目流。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0001 0001。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

3. reserved_bits

7位保留位，值恒为111 1111。

4. leak_valid_flag

1位标志位，定义了在内存缓冲区之间传输数据所采用的技术。

13.13.22 系统时钟描述符

这个MPEG-2描述符提供了一种说明关于系统时钟信息的方法，这个系统时钟用于生成时间戳。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

1. descriptor_tag

8位字段，值为0000 1011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0010。

3. external_clock_reference_indicator

1位的指示器，值为1表示系统时钟是由一个基准时钟所派生出来的，这个基准时钟可以在解码器中获得。

4. reserved_bit

保留位，值恒为1。

5. clock_accuracy_integer

6位的二进制数，与clock_accuracy_exponent联合使用，提供了系统时钟的分频准确性。

6. clock_accuracy_exponent

3位的二进制数，与clock_accuracy_integer联合使用，提供了系统时钟的分频准确性。

7. reserved_bits

5位保留位，值恒为1 1111。

13.13.23 目标背景网格描述符

686

这个MPEG-2描述符提供了在显示器指定位置内显示视频的方法。当视频不打算使用显示器全区域时，该描述符非常有用。对于DVB标准，当分辨率大于720×576（25Hz比特流）或720×480（30Hz比特流）时，需要用到这个描述符。

1. descriptor_tag

8位字段，值为0000 0111。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。

3. horizontal_size

14位的二进制数，定义了目标背景网格的水平大小（单位为像素）。

4. vertical_size

14位的二进制数，定义了目标背景网格的垂直大小（单位为像素）。

5. aspect_ratio_information

4位码字，定义了屏幕高宽比，如同在视频序列头中所定义的一样。

13.13.24　视频流描述符

这个MPEG-2描述符提供了用于标识一个视频基本流编码参数的基本信息。

1. descriptor_tag

8位字段，值为0000 0010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. multiple_frame_rate_flag

1位标志位，指示视频流中可能会出现多个帧速率。

4. frame_rate_code

4位码字，指示视频帧率，如表13-53所示。当multiple_frame_rate_flag为1时，这个特定帧率的指示同样允许视频流中出现其他帧率。

表13-53　frame_rate_code码字

指定的帧率	可能也包含这些帧率	编　码
禁止		0000
23.976		0001
24.0	23.976	0010
25.0		0011
29.97	23.976	0100
30.0	23.976、24.0、29.97	0101
50.0	25.0	0110
59.94	23.976、29.97	0111
60.0	23.976、24.0、29.97、30.0、59.94	1000
保留		1001
⋮		⋮
保留		1111

5. MPEG_1_only_flag

1位标志位，值为1表示视频流仅包含MPEG-1视频数据。

6. constrained_parameter_flag

如果MPEG_1_only_flag为0，这一位必须为1。如果MPEG_1_only_flag为1，这一位反映了在MPEG-1视频流中constrained_parameter_flag的值。

7. still_picture_flag

1位标志位，值为1表示视频流仅包含静态图像。值为0表示视频流可能不是静态图像就是动态图像。

8. profile_and_level_indication

可选的8位码字，反映了与MPEG-2视频流中的profile_and_level_indication字段所指示的档次和基本相同或者更高的档次和级别。仅当MPEG_1_only_flag＝0时这个字段才出现。

687

9. chroma_format

可选的两位码字，反映了与MPEG-2视频流中的chroma_format字段所指示的色度格式相同或者更高的色度格式。仅当MPEG_1_only_flag＝0时这个字段才出现。

10. frame_rate_extension_flag

1位可选标志位，值为1表示任何MPEG-2视频流中的frame_rate_extension_n和frame_rate_extension_d两个字段中任意一个是或者两者都是非零的。仅当MPEG_1_only_flag＝0时

这个字段才出现。

11. reserved_bits

5位可选的保留位，值恒为1 1111。仅当MPEG_1_only_flag＝0时这个字段才出现。

13.13.25 视频窗口描述符

这个MPEG-2描述符提供一种说明关于伴随视频基本流信息的方法。对于DVB标准，如果这个字段出现，解码器必须处理其数据。

1. descriptor_tag

8位字段，值为0000 1000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。

3. horizontal_offset

14位的二进制数，指示了视频窗口在目标背景网格上左上角像素的水平位置。 688

4. vertical_offset

14位的二进制数，指示了视频窗口在目标背景网格上左上角像素的垂直位置。

5. window_priority

4位二进制数，指示视频窗口如何重叠。值0000是最低优先级，值1111是最高优先级。高优先级窗口叠放在低优先级窗口之上。

13.14 MPEG-4 PMT/PSM描述

这些MPEG-4描述符用来标识出现在MPEG-2传输或者节目流中的MPEG-4相关的信息。它们通常包含在MPEG-2 PMT或PSM中的descriptor_loop中，也可能出现在其他MPEG-4相关的表中。

13.14.1 MPEG-4音频描述符

对于由PES包传送的独立的MPEG-4.3流，这个MPEG-4描述符提供了用于标识编码参数的基本信息。该描述符没有应用到封装在MPEG-4 SL包或FlexMux包中的MPEG-4.3流。

1. descriptor_tag

8位字段，值为0001 1100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

3. MPEG-4_audio_profile_and_level

8位字段，标识MPEG-4.3音频流的档次和级别。

13.14.2 外部ES ID描述符

这个MPEG-4描述符分配一个ED_ID（其定义在MPEG-4.1中给出）给一个节目元素，该节目元素没有ES_ID值。ES_ID允许指向一个场景描述符中的非MPEG-4组件或者关联一个IPMP流中的非MPEG-4组件。

对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。对于节目流，在PSM内，这个描述符是在elementary_stream_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0010 0000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0010。

3. external_ES_ID

16位字段，分配一个ES_ID（其定义在MPEG-4.1中给出）给一个节目的一个组件。

13.14.3　FMC描述符

这个MPEG-4描述符指示在封装到MPEG-2 PES包或MPEG-4节之前，MPEG-4 FlexMux工具已经被用于将MPEG-4 SL打包流复用到一个FlexMux流。该描述符将FlexMux信道关联到FluxMux流中SL打包流的ES_ID值。对于传输流中由elementary_PID值所指向的每一个节目元素，和节目流中传递FlexMux流的每一个elementary_stream_PID，都需要一个FMC描述符。

689

对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。对于节目流，在PSM内，这个描述符是在elementary_stream_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0001 1111。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. ES_ID

16位字段，定义了一个SL打包流的标识符。

4. flexMuxChannel

8位字段，定义了用于这个SL打包流的FlexMux信道数目。

13.14.4　FmxBufferSize描述符

这个MPEG-4描述符为每一个MPEG-4打包流传递FlexMux缓冲区大小的信息，而这些MPEG-4打包流将复用到MPEG-4 FlexMux流。每一个elementary_PID或者每一个传递FlexMux流的elementary_stream_PID都与一个FmxBufferSize描述符相关联。

对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。对于节目流，在PSM内，这个描述符是在elementary_stream_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0010 0010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. defaultFlexMuxBufferDescriptor()

这个描述符为FlexMux流指定了默认的FlexMux缓冲区大小。其定义在MPEG-4.1中给出了。

4. flexMux Buffer Descriptor()

这个描述符为FlexMux流中携带的SL打包流指定了FlexMux缓冲区大小。其定义在MPEG-4.1中给出了。

13.14.5　IOD描述符

这个MPEG-4描述符封装了MPEG-4.1 InitialObjectDescriptor结构。它使得可以通过标识MPEG-4.1场景描述和对象描述符流的ES_ID值来访问MPEG-4流。两者都包含关于作为场景一部分的MPEG-4流的进一步信息。

对于传输流，这个描述符是在PMT内的program_info_length之后的descriptor_loop中。对于节目流，在PSM内，这个描述符是在program_stream_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0001 1101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. scope_of_IOD_label

8位字段，定义了IOD_label字段的范围。值0x10表示IOD_label在节目流内或者传输流内的节目中是唯一的。值0x11表示IOD_label在携带IOD描述符的传输流内是唯一的。所有其他值都保留。 690

4. IOD_label

8位字段，定义了IOD描述符的标签。

5. initial object descriptor()

这个结构在MPEG-4.1中定义。

13.14.6 MultiplexBuffer描述符

这个MPEG-4描述符被关联到每一个包含一个MPEG-4 FlexMux或者SL打包流（包括那些包含MPEG-4部分的）的elementary_PID。

这个描述符仅应用到传输流，而没有应用到节目流。对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0010 0011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0110。

3. MB_buffer_size

24位字段，定义基本流的缓冲区[MB]的大小（单位是字节）。

4. TB_leak_rate

这个24位字段，为基本流定义了数据从传输缓冲区[TB]迁移到复用缓冲区[MB]的速率（单位是400bit/s）。

13.14.7 Muxcode描述符

这个MPEG-4描述符传递如MPEG-4.1所定义的MuxCodeTableEntry结构。MuxCodeTableEntry配置FlexMux的MuxCode模式。一个或多个Muxcode描述符可以与每一个elementary_PID或elementary_stream_ID相关联，从而利用MuxCode模式传送一个MPEG-4 FlexMux流。

对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。对于节目流，在PSM内，这个描述符是在elementary_stream_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0010 0001。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. MuxCodeTableEntry()

这个结构在MPEG-4.1中定义。

13.14.8 SL描述符

当单个MPEG-4 SL打包流被封装到MPEG-2 PES包中时，需要采用这个MPEG-4描述符。它将SL打包流的ES_ID与一个elementary_PID或elementary_stream_ID相关联。

这个描述符仅应用到传输流，而没有应用到节目流。对于传输流，这个描述符是在PMT内的ES_info_length之后的descriptor_loop中。

1. descriptor_tag

8位字段，值为0001 1110。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0010。

3. ES_ID

16位字段，定义了一个SL打包流的标识符。

13.14.9 MPEG-4视频描述符

对于由PES包传送的MPEG-4.2流，这个MPEG-4视频描述符提供了用于标识编码参数的基本信息。该描述符没有应用到封装成SL包或者FlexMux包的MPEG-4.2流。

1. descriptor_tag

691

8位字段，值为0001 1011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

3. MPEG-4_visual_profile_and_level

8位字段，标识MPEG-4.2视频流的档次和级别。它与伴随的MPEG-4.2流中的视觉对象序列头中的profile_and_level_indication有相同的值。

13.15 ARIB PMT描述符

这些ARIB描述符用于标识出现在MPEG-2传输流中的ARIB相关的信息。它们通常包含在MPEG-2 PMT的descriptor_loop中，也可能出现在其他ARIB相关的表中。ARIB描述符没有与PMT相关联，这一点将在第18章中讨论。

13.15.1 轮播兼容的复合描述符

这个ARIB描述符采用在数据轮播传送规范（ARIB STD-B24第3部分）中定义的描述符作为子描述符，通过应用子描述符的功能描述了累积控制。

13.15.2 组件描述符

这个ARIB描述符与在DVB描述符一节讨论的组件描述符是一样的。

13.15.3 条件播放描述符

这个ARIB描述符传送了条件播放的描述和发送ECM和EMM的PID。

13.15.4 内容可用性描述符

这个ARIB描述符描述了由接收者控制记录和输出内容的信息。Encryption_mode标志指示是

否加密数字视频输出。该描述符与数字复制控制描述符联合使用。

13.15.5 国家可用性描述符

这个ARIB描述符与在DVB描述符一节讨论的组件描述符是一样的。

692

13.15.6 数据组件描述符

这个ARIB描述符标识了数据编码系统标准。其语法与DVB的“数据广播ID描述符”是一致的（除字段名不相同以外）。

1. descriptor_tag

8位字段，值为1111 1101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. data_component_ID

16位字段，标识用于在广播网络中广播数据的数据广播规范。

注意：下面字段可能会重复[n]次。

4. additional_identifier_information

这个数据的定义取决于data_component_ID。

13.15.7 数字复制控制描述符

这个ARIB描述符包含了控制复制代次的信息。对于数字唱片，广播服务提供商采用该描述符来告知数字唱片设备有关事件记录和版权的信息。该描述符的descriptor_tag值为1100 0001。

13.15.8 紧急信息描述符

这个ARIB描述符被用来传送一个紧急事件警告信息。它可能仅被用于地面数字音频、地面数字电视、BS数字或宽带CS广播。该描述符同样也出现在ARIB网络信息表（Network Information Table，NIT）的descriptor_loop中。

1. descriptor_tag

8位字段，值为1111 1100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：下面7个字段可能会重复[n]次。

3. service_ID

16位字段，标识了广播节目编号。

4. start/end_flag

1位标志位，在紧急信息的起始处值为1，当传送结束时值为0。

5. signal_type

1位标志位，当传送类型1和类型2起始信号时，值分别为0和1。

6. reserved_bits

6位保留位，值恒为11 1111。

7. area_code_length
8位二进制数，指示这个字段之后的字节数。

注意：下面2个字段可能会重复[n]次。

8. area_code
12位字段，指示区域代码，如Notification No.45所定义的。
9. reserved_bits
693
4位保留位，值恒为1111。

13.15.9 层次传输描述符

这个ARIB描述符指示了当传输事件层次等级化时，各级流之间的关系。该描述符的descriptor_tag值为1100 0000。

13.15.10 链接描述符

这个ARIB描述符提供了一个到其他服务、传输流、节目指南、服务信息、软件更新等的链接。

13.15.11 Mosaic描述符

这个ARIB描述符与在DVB描述符一节讨论的Mosaic描述符是一样的。

13.15.12 家长分级控制描述符

这个ARIB描述符与在DVB描述符一节讨论的家长分级控制描述符是一样的。

13.15.13 流标识符描述符

这个ARIB描述符与在DVB描述符一节讨论的流标识符描述符是一样的。

13.15.14 系统管理描述符

这个ARIB描述符标识了广播的类型。它同样也出现在ARIB NIT的descriptor_loop中。
1. descriptor_tag
8位字段，值为1111 1110。
2. descriptor_length
8位二进制数，定义了这个字段之后的字节数。
3. system_management_ID
这个16位字段标识了广播的类型。
b0~b1：
00＝广播
01＝非广播
10＝非广播
11＝保留
b2-b7：
000000＝保留
000001＝CS数字广播

000010＝BS数字广播
000011＝地面数字TV广播
000100＝宽带CS数字广播
000101＝地面数字音频广播
000110-111111＝保留
剩余的8位（b8~b15）组成了additional_identifier_information字段，用于扩展广播信号标准。

13.15.15　目标区域描述符

这个ARIB描述符描述了目标区域，该目标区域属于一个事件或者包含一个事件的流的一部分。

13.15.16　视频解码控制描述符

当接收基于MPEG的静态图像时，这个ARIB描述符用来控制视频解码；当在视频结合点处编码格式发生改变时，这个ARIB描述符用于平滑显示视频。

1. descriptor_tag
8位字段，值为1100 1000。
2. descriptor_length
8位二进制数，定义了这个字段之后的字节数。

694

3. still_picture_flag
1位标志位，值为1表示1个静态（MPEG）图像；值为0表示动态图像。
4. sequence_end_code_flag
1位标志位，值为1表示视频流在序列结束时有序列结束码。值为0表示没有序列结束码。
5. video_encode_format
4位码字，详细说明了视频编码格式：
0000＝1080p
0001＝1080i
0010＝720p
0011＝480p
0100＝480i
0101＝240p
0110＝120p
0111~1111＝保留

13.16　ATSC PMT描述符

这些ATSC描述符用于标识出现在MPEG-2传输流中的ATSC相关的信息。它们通常包含在MPEG-2 PMT的descriptor_loop中，也可能出现在其他ARIB相关的表中。ATSC描述符没有与PMT相关联，这一点将在第15章中讨论。

13.16.1　AC-3音频流描述符

Dolby® Digital（AC-3）音频基本流可能被包含在MPEG-2比特流中，就如同标准MPEG音频流被包含在其中一样。类似于MPEG音频比特流，Dolby® Digital比特流被打包成PES包。

MPEG-2并不显性地支持Dolby® Digital比特流。同样，在PSI表中，MPEG-2音频流描述符并

没有充分地描述Dolby® Digital比特流的内容。

因此，包含Dolby® Digital或Dolby® Digital Plus音频数据的PES包将采用私有流1来发送。此外，还需要一个注册描述符Registration Descriptor（对于DVB系统是不需要的）和一个AC-3音频流描述符（descriptor_tag＝1000 0001）。

从2008年3月1日起，这个描述符必须携带3字节的ISO 639语言代码，并且如果存在，必须与ISO 639语言描述符携带的语言代码相匹配。

注意，对于ATSC和OpenCable™，这个AC-3音频描述符被命名为“AC-3音频流描述符”，而对于DVB，这个AC-3音频描述符被命名为“AC-3描述符”。这些描述符的语法在两个系统之间有明显的不同。

13.16.2　ATSC私有信息描述符

这个ATSC描述符提供了一种传送私有信息的途径。在单个descriptor_loop中可能会出现多个描述符。

1. descriptor_tag

8位字段，值为1010 1101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. format_identifier

32位的二进制数，定义了以下私有信息的所属者（这些所属者已经向SMPTE注册机构注册了），如表13-52所示。

695

4. private_data_byte

可选的[n]字节私有数据，由format_identifier所属者定义。

13.16.3　组件名称描述符

这个ATSC和OpenCable™描述符为服务的任何组件定义了一个可选的文本名称标签。

1. descriptor_tag

8位字段，值为1010 0011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. component_name_string()

名称字符串，基于ATSC的多字符串结构（Multiple String Structure）。

13.16.4　内容分级描述符

这个ATSC和OpenCable™描述符为给定节目定义等级。它同样也出现在ATSC和OpenCable™事件信息表（event information table，EIT）的descriptor_loop中。

1. descriptor_tag

8位字段，值为1000 0111。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. rating_region_count

6位二进制数，定义如下等级区域规范的次数（1~8）。

注意：[rating_region_count]定义了下面7个字段将重复的次数。

4. rating_region
8位二进制数，指定了下面数据所定义的等级区域。
5. rated_dimensions
8位二进制数，为内容分级指定了等级维度的编号，该内容分级是为节目所定义的。

注意：[rated_dimensions]定义了下面3个字段将重复的次数。

6. rating_dimension_j
8位二进制数，为由字段rating_region所定义的区域指定了在ATSC RRT实例中的维度索引。
7. reserved_bits
4位保留位，值恒为1111。
8. rating_value
4位二进制数，表示为由rating_region给定的区域所定义的维度（由字段rating_dimension_j所定义）的等级值。
9. rating_description_length
8位二进制数，定义了接下来的rating_description_text字段的长度（0~80）。
10.rating_description_text()
等级描述字符串，基于ATSC的多字符串结构（Multiple String Structure）。 696

13.16.5　增强信令描述符

这个ATSC描述符表示了节目元素的地面广播传输方法。如果节目元素与另一个节目元素是二选一的，该描述符将链接到可供选择的另一个节目元素，并且定义了广播台的优先选择。
1. descriptor_tag
8位字段，值为1011 0010。
2. descriptor_length
8位二进制数，定义了这个字段之后的字节数。
3. linkage_preference
2位码字，指示节目元素是否被链接到另一个节目元素。如果链接了，它同时也标识了广播台的优先选择。
00＝没有链接
01＝链接了，不优先选择
10＝链接了，首选
11＝链接了，不首选
4. tx_method
2位码字，定义了用来发送伴随节目元素的VSB发送方法。
00＝全速率，全编码
01＝半速率，1/2速率增强编码
10＝四分之一速率，1/4速率增强编码
11＝保留
5. linked_component_tag
可选的14位值，链接节目元素到另一个可供选择的节目元素。这个可供选择的节目元素是与

传输流中有相同lined_component_tag值的节目元素，该传输流的PMT与携带这个描述符的传输流的PMT有相同的program_number值。这个字段仅当linkage_preference＝01、10或11时才出现。

6. reserved_bits

可选的4位保留位，值恒为1111。仅当linkage_preference＝00时这个字段才出现。

13.16.6　再发行控制描述符

这个ATSC和OpenCable™描述符（也称作“广播标志”）传递由节目版权所有者所拥有的任何内容再发行控制信息。它同样也出现在ATSC和OpenCable™EIT的descriptor_loop中。

1. descriptor_tag

8位字段，值为1010 1010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. RC_information()

697

[n]字节可选的附加再发行控制信息，这些信息可以在将来再定义。

13.17　DVB PMT描述符

这些DVB描述符用于标识出现在MPEG-2传输流中的DVB相关的信息。它们通常包含在MPEG-2 PMT的descriptor_loop中，也可能出现在其他DVB相关的表中。DVB描述符没有与PMT相关联，这一点将在第17章中讨论。

13.17.1　AAC音频描述符

对于由PES包传送的MPEG-4 AAC、HE-AAC和HE-AAC v2音频流，这个DVB描述符提供了标识编码参数的基本信息。

1. descriptor_tag

8位字段，值为0111 1100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. profile_and_level

8位字段，标识了MPEG-4 AAC、HE-AAC或HE-AAC v2音频的档次和级别。

4. AAC_type_flag

1位标志位，表示AAC_type字段是否出现。

5. reserved_bits

7位保留位，值恒为000 0000。

6. AAC_type

可选的8位字段，指示MPEG-4 AAC、HE-AAC或HE-AAC v2基本流运载的音频的类型。仅当AAC_type_flag＝1时这个字段才出现。

7. additional_info

[n]字节可选信息。

13.17.2　AC-3和增强的AC-3描述符

Dolby® Digital（AC-3）或Dolby® Digital Plus（E-AC-3）音频基本流可能被包含在MPEG-2比

特流中，就如同标准MPEG音频流被包含在其中一样。类似于MPEG音频比特流，Dolby® Digital或Dolby® Digital Plus比特流被打包成PES包。

MPEG-2并不显性地支持Dolby® Digital或Dolby® Digital Plus比特流。同样，在PSI表中，MPEG-2音频流描述符并没有充分地描述Dolby® Digital或Dolby® Digital Plus比特流的内容。

因此，包含Dolby® Digital或Dolby® Digital Plus音频数据的PES包将采用私有流1来发送。此外，还需要一个AC-3描述符（descriptor_tag＝0110 1010），或者一个增强的AC-3描述符（descriptor_tag＝0111 1010）。

注意，对于ATSC和OpenCable™，这个AC-3音频描述符被命名为“AC-3音频流描述符”，而对于DVB，这个AC-3音频描述符被命名为“AC-3描述符”。这些描述符的语法在两个系统之间有明显的不同。

698

13.17.3 适配域数据描述符

这个DVB描述符用来在适配域的私有数据字段中指示所支持的数据字段类型。

1. descriptor_tag

8位字段，值为0111 0000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0x01。

3. adaptation_field_data_identifier

8位字段，标识在适配域的私有数据字段中传递的数据字段。某一位设置为1表示对应的数据字段是支持的。

b0＝声明交换数据字段
b1~b7＝保留

13.17.4 辅助数据描述符

这个DVB描述符用来指示在MPEG音频基本流中辅助数据是否出现，以及如果出现，其类型是什么。

1. descriptor_tag

8位字段，值为0110 1011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0x01。

3. ancillary_data_identifier

8位字段，标识在适配域的私有数据字段中传递的数据字段。某一位设置为1表示对应的数据字段是支持的。

b0＝DVD视频辅助数据
b1＝扩展的辅助数据
b2＝声明交换数据
b3＝DAB辅助数据
b4＝比例因子差错检查
b5＝保留
b6＝保留
b7＝保留

13.17.5 组件描述符

这个ARIB和DVB描述符指示流的类型，可以用来提供流的文本描述。对于DVB，该描述符仅出现在EIT和SIT中。

1. descriptor_tag

8位字段，值为0101 0000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. reserved_bits

4位保留位，值恒为0000。

4. stream_content

4位码字，指示流（音频、视频或数据）中内容的类型。

5. component_type

699

8位码字，指示音频、视频或数据的类型。

6. component_tag

8位字段，对于组件流，该字段与流标识描述符中的component_tag有相同的值。

7. ISO_639_language_code

24位字段，包含了一个3字符的语言代码。

8. text_char

[n]字节，定义流的文本描述。

13.17.6 国家可用性描述符

这个ARIB和DVB描述符标识允许或不允许接收服务的国家。对于每一个服务，这个描述符可能会出现两次，一次是列出允许接收服务的国家，第二次是列出不允许接收服务的国家。后一个列表覆盖前一个列表。

1. descriptor_tag

8位字段，值为0100 1001。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. country_availability_flag

1位标志位，值为1指示指定可接收服务国家的国家代码。值为0指示指定不可接收服务国家的国家代码。

4. country_code

[n]个24位字段，标识国家，采用3字符代码，如ISO 3166中所定义的。

13.17.7 数据广播ID描述符

这个DVB描述符标识了数据编码系统标准。

1. descriptor_tag

8位字段，值为0110 0100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. data_broadcast_ID

16位字段，标识用于在广播网络中广播数据的数据广播规范。这个字段的分配值，请参看ETR 162。

注意：下面的字段将重复[n]次。

4. ID_selector_byte

这个字段的定义取决于data_broadcast_ID。

13.17.8 DTS音频描述符

当DVB传输流中包含一个DTS®音频流时，这个描述符同样必须包括在这个DVB传输流中。其descriptor_tag值为0111 1011。

这个描述符或者MPEG-2注册描述符也必须出现在PMT和SIT中，以标识一个DTS®音频流。 700

13.17.9 扩展描述符

这个DVB描述符用来扩展descriptor_tag的8位值。其descriptor_tag值为0111 1111。

13.17.10 Mosaic描述符

mosaic组件是一个不同视频图像的集合，这个集合形成一个编码视频组件。这些图像的信息是有组织的，因此当每一个特定信息显示时，都是出现在屏幕上的一个小区域里。

这个ARIB和DVB描述符将一个数字视频组件分割成基本单元，分配这些基本单元到逻辑单元，并且将逻辑单元的内容与对应的信息（如特殊风格、服务、事件等）链接起来。该描述符的descriptor_tag值为0101 0001。

13.17.11 家长分级控制描述符

这个ARIB和DVB描述符给出了基于一定年龄而确定的等级，并且也提供了可能用于其他分级准则的扩展。对于DVB，该描述符仅出现在EIT和SIF中。

1. descriptor_tag

8位字段，值为0101 0101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：下面2个字段会重复[n]次。

3. country_code

24位字段，标识国家，采用3字符代码，如ISO 3166中所定义的。

4. rating

8位字段，指示建议的观看者最小年龄（单位为岁）。

13.17.12 私有数据说明符描述符

这个DVB描述符用来标识任何私有描述符或者描述符内私有字段的说明符。

1. descriptor_tag

8位字段，值为0101 1111。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0x04。

3. private_data_specifier

这个字段值的分配在ETR162中给出。

13.17.13 加扰描述符

这个DVB描述符指示了为加扰系统选择的操作模式。

1. descriptor_tag

8位字段，值为0110 0101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. scrambling_mode

701

8位值，标识了选择的DVB通用加扰算法（Common Scrambling Algorithm）模式。

13.17.14 服务切换描述符

这个DVB描述符使得当服务从一个传输流切换到另一个传输流时，解码器能够跟踪该服务。

1. descriptor_tag

8位字段，值为0110 0000。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. new_original_network_ID

16位的字段，定义了传输流的original_network_ID，在该传输流中服务在切换之后被发现了。

4. new_transport_stream_ID

16位的字段，定义了传输流的transport_network_ID，在该传输流中服务在切换之后被发现了。

5. new_service_ID

16位的字段，定义了切换之后服务的service_ID。

13.17.15 流标识符描述符

这个ARIB和DVB描述符使得特定的流与EIT中的一个描述相关联。该描述符用于一个服务中有多个相同类型流的情况。

1. descriptor_tag

8位字段，值为0101 0010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数，值为0x01。

3. component_tag

8位字段，标识了与组件描述符相关联的组件流。在PMT中，对于这个字段，每个流标识符描述符的值都不相同。

13.17.16 字幕描述符

这个DVB描述符标识了ETSI EN 300 743字幕数据。

1. descriptor_tag

8位字段，值为0101 1001。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：下面的4个字段可能重复[n]次，以允许单个描述符标识多个数据服务。

3. ISO_639_language_code

24位字段，包含了一个3字符的语言代码。

4. subtitling_type

8位字段，提供关于字幕内容和计划显示的信息。

5. composition_page_ID

16位字段，标识合成页。

702

6. ancillary_page_ID

16位字段，标识（可选的）补充页。

13.17.17 图文电视描述符

这个DVB描述符用来标识携带EBU图文电视数据的基本流。

1. descriptor_tag

8位字段，值为0101 0110。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：下面的4个字段可能重复[n]次，以允许单个描述符标识多个数据服务。

3. ISO_639_language_code

24位字段，包含一个3字符的语言代码。

4. teletext_type

5位码字，定义了图文电视页的类型：

0x01＝初始图文电视页

0x02＝图文电视字幕页

0x03＝附加信息页

0x04＝节目调度页

0x05＝为听力障碍人士提供的图文电视字幕页

5. teletext_magazine_number

3位二进制数，标识杂志编号。

6. teletext_page_number

8位字段，定义图文电视页编号（两个4位的16进制数字）。

13.17.18 VBI数据描述符

这个DVB和OpenCable™描述符定义了在伴随的打包基本流（PES）中的VBI服务类型。

1. descriptor_tag

8位字段，值为0100 0101。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

注意：下面的4个字段可能重复[n]次，以允许单个描述符标识多个数据服务。

3. data_service_ID

8位二进制数，标识出现在伴随基本流中的VBI数据的类型。其值如下：

0x01＝EBU图文电视

0x02＝带反转取景码的EBU图文电视

0x04＝视频编程系统（VPS）

0x05＝宽屏信令（WSS）

0x06＝隐藏字幕

0x07＝单色4：2：2采样点

0xF7＝场消隐期时间码（VITC）

0xF9＝复制保护

0xFB＝TV指南

0xFC＝NABTS

0xFE＝AMOL I/II

703

4. data_service_description_length

8位二进制数，指示当前字段之后的字节数。

注意：下面的字段将出现，并且当data_service_ID＝0x01、0x02、0x04、0x05、0x06、0x07、0xF7、0xFB、0xFC或0xFE时，这些字段可能重复[n]次。

5. reserved_bits

2位保留位，值恒为11。

6. field_parity

标志位，值为1表示场1数据，值为0表示场2数据。

7. line_offset

5位二进制数，定义了对于480i或576i视频信号，VBI数据被插入的行号。当field_parity＝0时，值263_D（480i）或313_D（576i）被添加到line_offset值，以获得行号。

注意：下面的字段将出现，并且当data_service_ID≠0x01、0x02、0x04、0x05、0x06、0x07、0xF7、0xFB、0xFC或0xFE时，这些字段可能重复[n]次。

8. reserved_bits

8位保留位，值恒为1111 1111。

13.17.19 VBI图文电视描述符

这个描述符的语法与图文电视描述符的语法是相同的。唯一的差别是它没有被用于将stream_type 0x06与VBI或者EBU图文电视标准相关联。解码器采用这个描述符中的语言选择杂志和字幕。该描述符的descriptor_tag值为0100 0110。

13.18 OpenCable PMT描述符

这些附加的描述符用于传递OpenCable™相关的信息。它们包含在MPEG-2 PMT的descriptor_loop中，并且也可能出现在其他OpenCable™相关的表中。OpenCable™描述符没有与PMT相关联，这一点将在第16章中讨论。

13.18.1 AC-3音频流描述符

这个OpenCable™描述符与在ATSC描述符一节中讨论的AC-3音频流描述符是一样的。

13.18.2 组件名称描述符

这个OpenCable™描述符与在ATSC描述符一节中讨论的组件名称描述符是一样的。

13.18.3 组件名称描述符

这个OpenCable™描述符为服务的任何组件都定义了一个可选的文本名称标签。

1. descriptor_tag

8位字段，值为1000 0100。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. reserved_bits

2位保留位，值恒为11。

4. string_count

6位二进制数，取值范围为1~63，定义了在这个描述符中定义的名称字符串的数量。 704

注意：[string_count]定义了下面3个字段重复的次数。

5. ISO_639_language

24位字段，包含一个3字符的语言代码。

6. string_length

8位二进制数，取值范围为1~31，定义了接下来的多语言名称字符串的长度（单位是字节）。

7. name_string()

可变长的文本字符串。

13.18.4 内容分级描述符

这个OpenCable™描述符与在ATSC描述符一节中讨论的内容分级描述符是一样的。

13.18.5 扩展视频描述符

这个OpenCable™描述符标识了可能需要处理的某些属性。

1. descriptor_tag

8位字段，值为1000 0011。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. catalog_mode_flag

1位标志位，值为1表示视频流支持从处理的比特流中选择和显示（视频停滞）单个帧的应用。值为0表示比特流正常地解码和显示。

4. video_includes_setup

1位标志位，值为1表示比特流中的视频包括一个7.5 IRE消隐脉冲电平。值为0表示比特流中的视频没有包括一个7.5 IRE消隐脉冲电平。

5. reserved_bits

6位保留位，值恒为11 1111。

13.18.6 帧率描述符

这个OpenCable™描述符标识了视频帧率。

1. descriptor_tag

8位字段，值为1000 0010。

2. descriptor_length

8位二进制数，定义了这个字段之后的字节数。

3. multiple_frame_rate_flag

1位标志位，值为1表示视频流中可能出现多个帧率。值为0表示视频流中仅出现一个单独的帧率。

4. frame_rate_code

4位码字，定义了出现在视频流中的帧率。如果multiple_frame_rate_flag是1，另外的帧率也可能会出现。

5. reserved_bits

705 3位保留位，值恒为111。

13.18.7 MAC地址列表描述符

当在MPEG-2传输流上实现IP多播时，采用这个OpenCable™描述符。当在PMT内通过分配stream_type＝0x0D而包含DSM-CC分段时，传送IP数据的流将被标识。

这个描述符用来标识由每一个数据基本流传送的数据（通过广播MAC组地址实现）。其descriptor_tag值为1010 1100。

13.18.8 再发行控制描述符

这个OpenCable™描述符与在ATSC描述符一节中讨论的再发行控制描述符是一样的。

13.18.9 VBI数据描述符

这个OpenCable™描述符与在DVB描述符一节中讨论的VBI数据描述符是一样的。

13.19 隐藏字幕

CEA-608和CEA-708是基本的隐藏字幕标准。而CEA-608（在第8章中已讨论）最初是设计与NTSC广播一起使用的，CEA-708是设计与数字TV广播一起使用的。

当前还没有用于采用HDMI或者480p/720p/1080i/1080p模拟YPbPr接口传递隐藏字幕数据到TV的标准。在这些情况下，隐藏字幕解码和显示必须在解码压缩视频的机顶盒中完成。

13.19.1 CEA-708

CEA-708 DTV隐藏字幕标准对基于NTSC的CEA-608隐藏字幕标准进行了许多改变。焦点集中在给浏览者更好的观看信息，给他们更多的控制权。

最重要的是，CEA-708中可以包括更多的信息。对于字幕信息，CEA-608支持最多960bit/s，而CEA-708为字幕信息保留一个9600bit/s的恒定带宽（包括用于CEA-608字幕的960bit/s）。

浏览者可以控制字幕文本的大小。那些视力较差的人可以将字幕放大，那些更喜欢字幕不要覆盖图像较大区域的人可以将字幕缩小，其他人也可以不管它。

CEA-708同样提供了更多的字母和符号，支持多语言字幕。而CEA-608字符集没有某些语言如法语、西班牙语、德语、意大利语或葡萄牙语字幕所需的所有字母和口音标记，CEA-708提供了这些语言所需的全部，并且提供了更多的语言字幕内容。

CEA-708支持多字体和更多颜色，消除了常见的沉闷的等宽黑底白字视觉，定义了8种字体（包括均匀间距、非正式的和脚本字体）和最多64种文本及背景颜色，尽管字幕解码器不需要支持所有的字体和颜色。这使得字幕设计者能够便利地改善字幕视觉效果。然而，他们将不得不考虑在不需要多字体支持的情况下，字幕将如何显示在电视上。

附加的颜色支持意味着传统的黑盒背景可以被一个彩色盒替代，完全有利于废除有边的或下拉阴影的文本。字幕盒也可以做成半透明的（透明的）。 706

CEA-708允许利用视频比特流的序列、GOP或图像层的user_data，从而添加隐藏字幕数据到MPEG-2传输流。图13-25简要说明了DTV隐藏字幕协议模型。

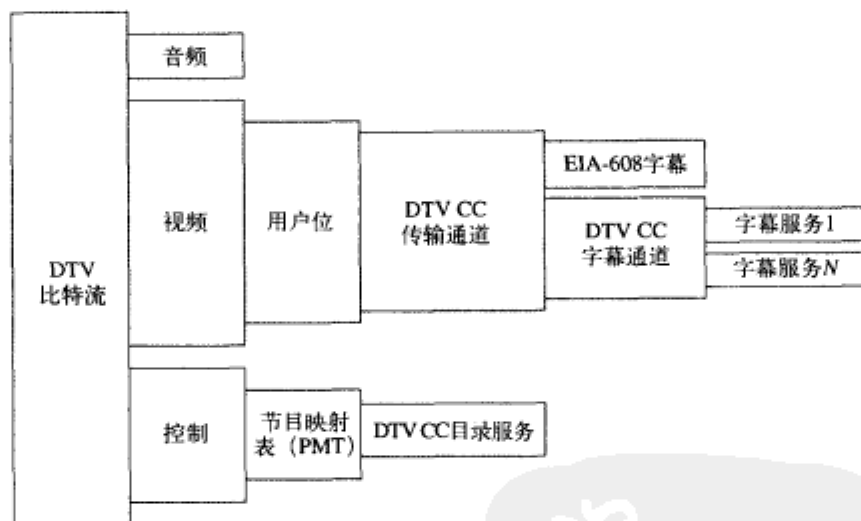


图13-25 MPEG-2比特流中DTV隐藏字幕数据

CEA-608字幕数据没有嵌入到DTV协议栈。这使得它可以在传输层抽取出来，从而可以设计更小的字幕解码器，因为不需要解析整个DTV隐藏字幕信道比特流来获得两字节的CEA-608数据。

MPEG编码图像的传输顺序与显示顺序是不相同的。字幕数据采用类似的方法重排序，因此也必须与其对应的图像一起重排序（由解码器执行），这个过程先于包定位和抽取。

13.19.2 MPEG-2视频

CEA-708隐藏字幕采用从信号带宽中分配的连续9600bit/s流。字幕数据是一帧一帧分配的，因此每秒传输的数据为1200字节。这使得对于480i30或1080i30信号，每帧最多有20字节的字幕数据。平均起来，CEA-608字幕分配速率为960bit/s，而CEA-708字幕分配速率为8640bit/s。

隐藏字幕可能仅出现在图像层。比特流语法如下。

1. user_data_start_code

32位字符串，值为0x000001B2。

2. user_identifier

32位标识符，值为0x47413934，表示用户数据遵循CEA-708隐藏字幕规范。

3. user_data_type_code

8位码字，其值定义了接下来信息的类型，如表13-54所示。其他值不是用于其他标准，就是保留以备将来使用。

表13-54 user_data_type_code码字

数据类型	编 码	ATSC	DVB	OpenCable	SCTE21
隐藏字幕	0x03	是	是	是	是
附加的CEA-608 数据	0x04	–	–	是	是
亮度PAM数据	0x05	–	–	是	是
彩条信息	0x06	是	是	是	是

注意：当user_data_type_code＝0x03时，下面7个字段将出现。

4. reserved_bit

保留位，值恒为1。

5. process_cc_data_flag

1位标志位，如果设置为1，cc_data必须处理。如果设置为0，cc_data必须丢弃。

6. zero_bit

零位，值恒为0。

7. cc_count

5位二进制数，取值范围为0~31，指示跟随当前字段之后的隐藏字幕结构的数目。设定该字段的值，从而为隐藏字幕数据维持一个9600bit/s的固定带宽。

8. reserved_bits

8位保留位，值恒为1111 1111。

注意：[cc_count]定义了下面5个字段会重复的次数。

707~708

9. marker_bits

5位标记位，值恒为1 1111。

10. cc_valid

如果这一位的值为1，两个隐藏字幕数据字节是有效的。如果设置为0，两个数据字节是无效的。无效表示CEA-608的时钟插补（run-in）和起始位都没有生成。

11. cc_type

2位码字，指定了接下来的隐藏字幕数据的类型，如表13-55所示。

表13-55 cc_type码字

CC类型	编码
CEA-608行21场1字幕	00
CEA-608行21场2字幕	01
CEA-708通道包数据	10
CEA-708通道包起始	11

12. cc_data_1

隐藏字幕数据的第一个8位。仅当process_cc_data_flag为1时才处理这些数据。

13. cc_data_2

隐藏字幕数据的第二个8位。仅当process_cc_data_flag为1时才处理这些数据。

14. marker_bits

8位标记位，值恒为1111 1111。

注意：当user_data_type_code＝0x04时，下面7个字段会出现。

15. marker_bits

3位标记位，值恒为111。

16. additional_cc_count

5位二进制数，定义了CEA-608数据出现的行数。

注意：[additional_cc_count]定义了下面5个字段重复的次数。

17. additional_cc_vaild

如果这一位设置为1，两个隐藏字幕数据字节是有效的。如果设置为0，这两个数据字节是无效的。

18. additional_cc_line_offset

5位二进制数，定义了CEA-608隐藏字幕数据起始行相对于行9和行272（对于480i系统），或者相对于行5和行318（对于576i系统）的行偏移量。

19. additional_cc_field_number

2位码字，指示CEA-608数据开始的场的编号（以显示顺序为序）。

00＝禁止

01＝第1场

10＝第2场

11＝第3个显示场（在电影模式中为重复场）

20. additional_cc_data_1

CEA-608隐藏字幕数据的第一个8位。仅当additional_cc_valid为1时才处理这些数据。

21. additional_cc_data_2

CEA-608隐藏字幕数据的第二个8位。仅当additional_cc_valid为1时才处理这些数据。 709

注意：当采用一个没有列在表13-54中的user_data_type_code时，下面字段会出现。

22. reserved_user_data

可选的8位保留位。

13.19.3 MPEG-2视频（SCTE 21）

被OpenCable™所使用的SCTE 21规范，为有线数字电视隐藏字幕构成了基础。

SCTE 21扩展了ATSC隐藏字幕，以更好地支持CEA-608字幕，这是因为某些有线电视系统在其他VBI线上采用CEA-608隐藏字幕格式传送非字幕数据。

SCTE 21中可以采用一个脉幅调制（Pulse Amplitude Modulation，PAM）机制来传送基本的VBI波形，如：

- 为除行21以外的一个或多个VBI行采用的CEA-608兼容的隐藏字幕
- SID/AMOL信号
- 基于EIA-516 NABTS规范的北美基本图文电视
- 世界图文电视系统（WST）

- 场消隐期时间码（VITC）

尽管大多数标准采用两级亮度编码，但是也支持多级PAM编码。

13.19.4 MPEG-2视频（SCTE 20）

SCTE 20规范是有线数字电视隐藏字幕的早期标准，它有可能作为user_data出现的图像层。OpenCable™和其他有线数字电视解码器都必须支持SCTE 20和21，以确保与当前系统的向后兼容性。

1. user_data_start_code

32位字符串，值为0x000001B2。

2. user_data_type_code

8位码字，值为0x03，指示隐藏字幕信息。

3. reserved_bit

7位保留位，值恒为100 0000。然而，某些早期的有线电视系统替代采用000 0000，因此，解码器应该忽略该字段的值。

4. VBI_data_flag

如果这一位设置为1，接下来将有一个或多个VBI数据结构。

5. cc_count

710

5位二进制数，取值范围为0~31，指示当前字段之后的隐藏字幕结构的数目。

注意：[cc_count]定义了下面6个字段重复的次数。

6. cc_priority

2位码字，表示结构在图像重构中的优先级。00为最高优先级，11为最低优先级。

7. field_number

2位码字，指示CEA-608数据开始的场的编号（以显示顺序为序）。

00＝禁止

01＝第1场

10＝第2场

11＝第3个显示场（在电影模式中为重复场）

8. line_offset

5位二进制数，定义了CEA-608隐藏字幕数据起始行相对于行10和行273（对于480i系统），或者相对于行6和行319（对于576i系统）的行偏移量。

9. cc_data_1

隐藏字幕数据的第一个8位。

10. cc_data_2

隐藏字幕数据的第二个8位。

11. marker_bit

标记位，值恒为1。

12. non_real_time_video_count

4位二进制数，取值范围为0~15，表示接下来的非实时视频结构的数目。

注意：[non_real_time_video_count]定义了下面8个字段重复的次数。

13. non_real_time_video_priority

2位码字，表示结构在非实时VBI数据中的优先级。00为最高优先级，11为最低优先级。

14. sequence_number

2位二进制数，随着序列增加而增加（增量为1）。值为00表示非实时采样视频行将不会被重构（待用），直到收到一个非零的sequence_number的段。

15. non_real_time_video_field_number

1位标志位，表示是否重构数据到奇数场（0）或偶数场（1）。

16. line_offset

5位二进制数，定义了VBI数据起始行相对于行10和行273（对于480i系统），或者相对于行6和行319（对于576i系统）的行偏移量。

注意：当sequence_number≠00时，下面4个字段会出现。

17. segment_number

5位二进制数，定义了非实时采样视频段的编号，从0 0001开始。

非实时采样视频被分割成64字节的段，每一个段都作为一个32亮度（Y）采样点数组跟一个16色度采样点数组对（Cb，Cr）传输，起始位是最左边像素的最高有效位。在任何同一非实时视频行的新像素段之前，视频序列的所有段都应该按顺序传输。 711

注意：下面字段会重复32次。

18. non_real_time_video_Y_data

这个段的8位非实时Y数据。

注意：下面2个字段将重复16次。

19. non_real_time_video_Cb_data

这个段的8位非实时Cb数据。

20. non_real_time_video_Cr_data

这个段的8位非实时Cr数据。

13.19.5 MPEG-4.10（H.264）视频

MPEG-4.10（H.264）视频的隐藏字幕采用视频基本流的SEI RBSP语法来传输。

MPEG-2字幕语法的user_data_start_code字段被下面字段所替代。

1. itu_t_t35_country_code

8位字段，值为0xB5。

2. itu_t_t35_provider_code

16位字段，值为0x0031。

13.19.6 SMPTE 421M（VC-1）视频

SMPTE 421M（VC-1）视频的隐藏字幕是可选的，采用视频基本流的用户数据来传输。

MPEG-2字幕语法的user_data_start_code字段被下面字段所替代。

VCI_user_data_start_code

32位字符串，值为0x0000011D，表示user_data的开始。

13.19.7 VBI标准

下一节将要讨论的“VBI标准”还定义了如何添加隐藏字幕数据到MPEG-2传输流。

13.20 VBI标准

ETSI EN 301 775和OpenCable™标准为DVB和有线数字电视应用定义了如何添加隐藏字幕、图文电视、视频编程系统（VPS）、宽屏信令（WSS）数据等到一个MPEG-2传输流。MPEG-2 PES包以私有流1的形式传送数据，而这些PES包依次由传输包传送。尽管最初是为DVB标准所设计的，但是它可以应用到任何MPEG-2比特流。在VBI标准中需要用到DVB VBI图文电视描述符。

PES数据字段的语法如下。

13.20.1 data_identifier

8位二进制数，标识了在PES包中传输的数据类型。其取值范围为0x10~0x1F以及0x99~0x9B。对于OpenCable™，其值为0x99。

712

注意：下面字段可能重复[n]次，以使得在单个流内传输多个类型的数据。

13.20.2 data_unit_ID

8位二进制数，标识了出现的数据类型。其值如下：

0x02＝EBU图文电视非字幕数据
0x03＝EBU图文电视字幕数据
0xC0＝带反转取景码的EBU图文电视
0xC3＝视频编程系统（VPS）
0xC4＝宽屏信令（WSS）
0xC5＝隐藏字幕
0xC6＝单色4：2：2采样点
0xD0＝AMOL I
0xD1＝AMOL II
0xD5＝美国EBU图文电视（NABTS）
0xD6＝TV指南
0xD7＝复制保护
0xD9＝场消隐期时间码（VITC）
0xFF＝填充

13.20.3 data_unit_length

8位二进制数，标识这个字段之后的字节数。如果有一个在0x10和0x1F之间的值包含在data_identifier中，这个字段必须设置为0x2C。

注意：当data_unit_ID＝0x02、0x03或0xC0时，下面的字段将出现。这个包用于传递EBU图文电视信息（ETSI EN 300 706）。

13.20.4 reserved_bits

2位保留位，值恒为11。

13.20.5 field_parity

1位标志位，值为1表示场1数据，值为0表示场2数据。

13.20.6 line_offset

5位二进制数，定义了576i视频中图文电视数据插入的行号。当field_parity＝0时，值313_D被添加到line_offset值。

13.20.7 framing_code

8位字段，定义了所采用的取景代码。对于EBU图文电视，其值为1110 0100。对于EBU逆向图文电视，其值为0001 1011。

逆向图文电视用于传送不计划让公众接收的信号，如为有线电视初始设备和计划表等所使用的下行流控制信号。逆向图文电视的使用正在走下坡路，许多广播台现在采用图文电视包31来取而代之。

13.20.8 txt_data_block

336位的字段，对应跟随在时钟插补（时钟同步）和取景代码（字节同步）之后的42字节576i EBU图文电视数据。

注意：当data_unit_ID＝0xC3时，下列字段将出现。这个包用于传递VPS信息（ETSI EN 300 231）。

13.20.9 reserved_bits

2位保留位，值恒为11。

13.20.10 field_parity

当这一位设置为1时，解码器仅需要执行field_parity，当设置为0时，解码器可以将包忽略。 713

13.20.11 line_offset

5位二进制数，为576i视频信号定义了VPS数据插入的行号。当这个字段值为1 0000时，解码器仅需要执行line_offset行，可以忽略其他行。

13.20.12 VPS_data_block

104位字段，对应跟随在时钟插补和起始码数据之后的13字节576i VPS数据。

注意：当data_unit_ID＝0xC4时，下面的字段将出现。这个包用于传递WSS信息（ETSI EN 300 294）。

13.20.13 reserved_bits

2位保留位，值恒为11。

13.20.14 field_parity

当这一位设置为1时，解码器仅需要执行field_parity，当设置为0时，解码器可以将包忽略。

13.20.15 line_offset

5位二进制数，为576i视频信号定义了WSS数据插入的行号。当这个字段值为1 0111时，解码器仅需要执行line_offset行，可以忽略其他行。

13.20.16 WSS_data_block

14位字段，对应跟随在时钟插补（run-in）和起始码数据之后的14位576i WSS数据。

13.20.17 reserved_bits

2位保留位，值恒为11。

注意：当data_unit_ID＝0xC5时，下面的字段将出现。这个包用于传递CEA-608隐藏字幕信息。

13.20.18 reserved_bits

2位保留位，值恒为11。

13.20.19 field_parity

714 1位标志位，值为1表示场1（行21）数据，值为0表示场2（行284）数据。

13.20.20 line_offset

5位二进制数，为480i视频信号定义了字幕数据插入的行号。当这个字段值为1 0101时，解码器仅需要执行line_offset行，可以忽略其他行。

13.20.21 closed_captioning_data_block

16位字段，对应16位的480i CEA-608隐藏字幕数据。

注意：当data_unit_ID＝0xD0时，下面的字段将出现。这个包用于传递AMOL I信息。

13.20.22 reserved_bits

2位保留位，值恒为11。

13.20.23 field_parity

1位标志位，值为1表示场1数据，值为0表示场2数据。

13.20.24 line_offset

5位二进制数，为480i视频信号定义了AMOL I数据插入的行号。有效值为0 1010~10110（10_D~22_D）。

13.20.25 AMOL48_data_block

41位字段，对应跟在7位AMOL I start_of_message头之后的41位480i AMOL I数据。

13.20.26 trailer

7位字段，值恒为000 0000。

注意：当data_unit_ID＝0xD1时，下面的字段将出现。这个包用于传递AMOL II信息。

13.20.27 reserved_bits

2位保留位，值恒为11。

13.20.28 field_parity

1位标志位，值为1表示场1数据；值为0表示场2数据。

13.20.29 line_offset

5位二进制数，为480i视频信号定义了AMOL II数据插入的行号。有效值为0 1010~1 0110（10_D~22_D）。

13.20.30 AMOL96_data_block

88位字段，对应跟在8位AMOL II start_of_message头之后的88位480i AMOL II数据。

注意：当data_unit_ID＝0xD5时，下面的字段将出现。这个包用于传递美国图文电视（NABTS）信息。

13.20.31 reserved_bits

2位保留位，值恒为11。

13.20.32 field_parity

1位标志位，值为1表示场1数据；值为0表示场2数据。

13.20.33 line_offset

5位二进制数，为480i视频信号定义了NABTS数据插入的行号。有效值为0 1010~1 0110（10_D~22_D）。

13.20.34 framing_code

8位二进制数，值为1110 0111。

13.20.35 NABTS_data_block

264位字段，对应跟随在时钟插补（时钟同步）和取景代码（字节同步）之后的33字节NABTS数据。 715

注意：当data_unit_ID＝0xD6时，下面的字段将出现。这个包用于传递TV指南信息。

13.20.36 reserved_bits

2位保留位，值恒为11。

13.20.37 field_parity

1位标志位，值为1表示场1数据，值为0表示场2数据。

13.20.38 line_offset

5位二进制数，为480i视频信号定义了TV指南数据插入的行号。有效值为0 1010~1 0110（10_D~22_D）。

13.20.39 TVG2X_data_block

32位字段，对应跟随在时钟插补和取景代码之后的32位480i TV指南数据。

注意：当data_unit_ID＝0xD7时，下面的字段将出现。这个包用于传递复制保护信息。

13.20.40 reserved_bits

2位保留位，值恒为11。

13.20.41 field_parity

1位标志位，值为1表示场1数据，值为0表示场2数据。

13.20.42 line_offset

5位二进制数，为480i视频信号定义的复制保护数据插入行号。有效值为0 1100~1 0110（14_D~22_D）。

13.20.43 CP_data_block

2位字段，对应IEC 61880 B.2分段的位7和位8。

13.20.44 reserved_bits

6位保留位，值恒为11 1111。

注意：当data_unit_ID＝0xD9时，下面的字段将出现。这个包用于传递VITC信息。

13.20.45 reserved_bits

2位保留位，值恒为11。

13.20.46 field_parity

1位标志位，值为1表示场1数据，值为0表示场2数据。

13.20.47 line_offset

5位二进制数，为480i视频信号定义了VITC数据插入的行号。有效值为0 1100~1 0110 (14_D~22_D)。

13.20.48 VITC_data_block

64位字段，对应VITC的64个有效数据位（除同步位之外）。

716

注意：当data_unit_ID＝0xFF时，下面的字段将出现。

13.20.49 stuffing_type

[n]字节字段，每字节值为1111 1111，任何数量的填充字节都可能出现，如果出现，解码器将其忽略。

13.21 图文电视

目前有两个通过MPEG-2传输图文电视的标准，DVB的“VBI标准”（参见前一节）和新近的“EBU图文电视”标准。“EBU图文电视”只支持EBU图文电视数据的传输。DVB的“VBI标准”允许传输多种类型的VBI数据。

某些系统可能只支持“EBU图文电视”标准，而其他系统可能只支持DVB的“VBI标准”，还有一些两大标准都支持。在两大标准都必须支持的系统中，两大标准分别采用各自的PID；图文电视数据在两个PID上都需要广播。

DVB EBU图文电视标准

对于DVB应用，ETSI EN 300 472标准定义了如何将EBU图文电视添加到一个MPEG-2传输流中。MPEG-2 PES包中以私有流1形式传送数据，而这些PES包依次由传输包传送。尽管为DVB标准设计，但对于任意MPEG-2比特流，它同样适用。要求使用DVB图文电视描述符。

PES数据字段的对应语法如下。

1. data_identifier

8位二进制数，标识PES包中传送的数据类型。值为0x10~0x1F。

注意：下列字段可能会重复[n]次。

2. data_unit_ID

8位二进制数，标识出现的数据类型。值为：0x02＝EBU图文电视非字幕数据，0x03＝EBU图文电视字幕数据。

3. data_unit_length

8位二进制数，指示这一字段之后的字节数，值为0x2C。

4. reserved_bits

2位保留位，值为11。

5. field_parity

1位标记位，值为1指示场1数据，值为0指示场2数据。

6. line_offset

5位二进制数，定义了对于一个576i 视频信号，图文电视数据插入的行数。该字段的值只有

在0 0111和1 0110之间是有效的。当field_parity＝0时，line_offset的值加上313_D得到行数。

7. framing_code

717

8位字段，定义了采用的帧编码。值为1110 0100。

8. magazine_and_packet_address

16位字段，对应杂志和包地址。

9. data_block

320位字段，对应图文电视数据保留的40字节。

13.22 有效格式描述（AFD）

一旦包含有用信息的矩形图像无法扩展到编码帧的全长或全宽，视频用户数据应包括AFD（ATSC A/53、ETSI TS 101154以及CEA-805的一部分）。AFD的功能与第8章中介绍的宽屏信令（WSS）类似。

13.22.1 MPEG-2视频

在序列扩展、GOP头和图像编码扩展之后，AFD在视频基本比特流的用户数据中是可选传送的。

1. user_data_start_code

32位字符串，值为0x000001B2，指示user_data的开始。

2. user_identifier

32位标识位，值为0x44544731，指示AFD语法。

3. zero_bit

值恒为0。

4. active_format_flag

1位标志位，如果值设为1，在该数据结构中描述有效格式。

5. reserved_bits

保留位，值恒为00 0001。

6. reserved_bits

可选保留位，值恒为1111。当active_format_flag＝1时该字段才出现。

7. active_format

可选字段，如表13-56所示，定义了感兴趣的区域，只有当active_format_flag＝1时该字段才出现。

13.22.2 MPEG-4.10（H.264）视频

在视频基本流的SEI RBSP语法中，AFD是可选的传送。

MPEG-2 AFD语法的user_data_start_code字段由下列两个字段替代。

1. itu_t_t35_country_code

8位字段，值为0xB5。

2. itu_t_t35_provider_code

16位字段，值为0x0031。

13.22.3 SMPTE 421M（VC-1）视频

在视频基本流的用户数据中，AFD是可选传送的。

MPEG-2 AFD语法的user_data_start_code字段被下面的字段所代替。

VC1_user_data_start_code

32位字符串，值为0x0000011D，指示user_data的开始。

718

表13-56 AFD active_format值。AFD与条数据的任意组合可能在视频用户数据中出现（两者选一、都不出现或都出现）。需要注意的是，由于AFD仅处理4：3、14：9和16：9的宽高比，因此AFD数据可能无法一直与条数据完全匹配，然而条数据能够正确代表几乎所有的宽高比。如果AFD数据与条数据冲突，那么除非AFD＝0000或0100，否则AFD数据将取得优先权

AFD Active_Format Value（R3~R0）	WSS位 b3、b2、b1、b0	4：3编码帧	16：9编码帧
0000~0001	保留		
0010	0100	不推荐	不推荐
0011	0010	不推荐	不推荐
0100	1101	大于16：9的宽高比	大于16：9的宽高比
0101~0111	保留		
1000	–	4：3全帧图像	16：9全帧图像
1001	1000	4：3全帧图像	4：3邮筒图像
1010	1011	16：9 Letterbox图像	16：9全帧图像
1011	0001	16：9 Letterbox图像	14：9邮筒图像
1100	保留		
1101	1110	4：3全帧图像，备选14：9中心	4：3邮筒图像，备选16：9中心
1110	–	16：9 Letterbox图像，备选14：9中心	16：9全帧图像，备选4：3中心
1111	–	16：9 Letterbox图像，备选4：3中心	16：9全帧图像，备选4：3中心

719

13.23 字幕

字幕由一个或多个压缩的位图图像以及每个图像对应的可选矩形背景组成。字幕被定位于给定的位置。对于指定数目的视频帧，字幕从定义好的开始时间开始显示。

位图技术支持任何语言，而不仅仅是那些解码器支持的语言，而且使字幕制作者能够完全控制字符的显示，包括字体大小和间距。此外，还易于使用不是任何标准字符集一部分的字符和符号，例如在代表特定名称的表意语言中出现的那些字符。

13.23.1 数字有线电视字幕

SCTE 27定义中的subtitle_massage()定义了节目相关的字幕位图。参考节目的节目时钟（PCR），字幕文本的显示以显示时间标识（PTS）的形式给出。

subtitle_massage()在传输流包中传送，其PID＝0x1FFB。语法如下。

1. table_ID

8位码字，值为0xC6。

2. reserved_bits

2位保留位，值恒为00。

3. reserved_bits

2位保留位，值恒为11。

4. section_length

12位二进制数，定义了该字段后的字节数，等于并包括CRC_32。

5. reserved_bit

1位保留位，值恒为0。

6. segmentation_overlay_included

1位标志位，值为1表示消息包括分割定义。

7. protocol_version

6位二进制数，将来允许消息类型传送参数，这些参数的构造可能不同于那些当前协议定义的参数。当前，字幕消息仅定义protocol_version值为00 0000。

8. table_extension

16位二进制数，用来区分不同类型的已分割message_body()，这些message_body()在传输流中同时出现，并全部采用subtitle_message()传送。当segmentation_overlay_included＝1时该字段才出现。

9. last_segment_number

12位二进制数，指示恢复完整消息所需的最后片断的段数。当segmentation_overlay_included＝1时该字段才出现。

720

10. segment_number

12位二进制数，指示（可能的）多部分消息中哪个部分将出现。当segmentation_overlay_included＝1时该字段才出现。

11. ISO_639_language_code

24位字段，包含3字符语言编码。

12. pre_clear_display

标志位，值为1表示在字幕文本显示之前，整个显示都将变为透明的。否则，字幕文本将加入到屏幕上已存在的文字中。

13. immediate

1位标志位，值为1表示字幕收到后将立即显示。否则将提示在display_in_PTS时间显示。

14. reserved_bit

保留位，值恒为0。

15. display_standard

5位码字，定义了字幕准备显示的格式。

00000＝720×480

00001＝720×576

00010＝1280×720

00011＝1920×1080

16. display_in_PTS

当这个32位值与33位MPEG节目时钟（90kHz）的32最低有效位匹配时，将显示字幕。

17. subtitle_type

4位码字，指示字幕数据时钟的格式。当前仅定义了一个值0001。

18. reserved_bit

保留位，值恒为0。

19. display_duration

11位二进制数，指示字幕数据显示的视频帧数目，取值范围为1~2000。

20. block_length

16位二进制数，指示之后出现的字节数，不包括CRC和任何描述符。

注意：当subtitle_type＝0001时出现下列字段。

21. reserved_bits

5位保留位，值恒为0 0000。

22. background_style

1位标志位，定义了背景类型：0＝透明，1＝帧。

23. outline_style

2位码字，定义了文本轮廓类型：00＝无，01＝轮廓，10＝下拉阴影，11＝保留。

24. bitmap_Y_component

5位二进制数，为文本颜色定义了Y值，取值范围为0~31。

25. bitmap_opaque_enable

1位标志位，值为1表示文本颜色应为不透明的（无视频混合）。值为0表示将与视频做50%的混合。

721

26. bitmap_Cr_component

5位二进制数，为文本颜色定义了Cr值，取值范围为0~31。

27. bitmap_Cb_component

5位二进制数，为文本颜色定义了Cb值，取值范围为0~31。

28. bitmap_top_H_coordinate

12位二进制数，取值范围为0~1919，定义了解压位图最左边像素的水平坐标。

29. bitmap_top_V_coordinate

12位二进制数，取值范围为0~1079，定义了解压位图顶行的垂直坐标。

30. bitmap_bottom_H_coordinate

12位二进制数，取值范围为0~1919，定义了解压位图最右边像素的水平坐标。

31. bitmap_bottom_V_coordinate

12位二进制数，取值范围为0~1079，定义了解压位图底行的垂直坐标。

32. frame_top_H_coordinate

12位二进制数，取值范围为0~1919，定义了帧最左边像素的水平坐标。当background_style＝1时该字段才出现。

33. frame_top_V_coordinate

12位二进制数，取值范围为0~1079，定义了帧顶行的垂直坐标。当background_style＝1时该字段才出现。

34. frame_bottom_H_coordinate

12位二进制数，取值范围为0~1919，定义了帧最右边像素的水平坐标。当background_style＝1时该字段才出现。

35. frame_bottom_V_coordinate

12位二进制数，取值范围为0~1079，定义了帧底行的垂直坐标。当background_style＝1时该字段才出现。

36. frame_Y_component

5位二进制数，为帧颜色定义了Y值，取值范围为0~31。当background_style＝1时该字段才出

现。

37. frame_opaque_enable

1位标志位，值为1表示帧颜色应为不透明的（无视频混合）。值为0表示将与视频做50%的混合。当background_style＝1时该字段才出现。

38. frame_Cr_component

5位二进制数，为帧颜色定义了Cr值，取值范围为0~31。当background_style＝1时该字段才出

722 现。

39. frame_Cb_component

5位二进制数，为帧颜色定义了Cb值，取值范围为0~31。当background_style＝1时该字段才出现。

40. reserved_bits

4位保留位，值恒为0000。当outline_style＝01时该字段才出现。

41. outline_thickness

4位二进制数，取值范围为0~31，定义了文本轮廓粗细度。当outline_style＝01时该字段才出现。

42. outline_Y_component

5位二进制数，为文本轮廓颜色定义了Y值，取值范围为0~31。当outline_style＝01时该字段才出现。

43. outline_opaque_enable

1位标志位，值为1表示文本轮廓颜色应为不透明的（无视频混合）。值为0表示将与视频做50%的混合。当outline_style＝01时该字段才出现。

44. outline_Cr_component

5位二进制数，为文本轮廓颜色定义了Cr值，取值范围为0~31。当outline_style＝01时该字段才出现。

45. outline_Cb_component

5位二进制数，为文本轮廓颜色定义了Cb值，取值范围为0~31。当outline_style＝01时该字段才出现。

46. shadow_right

4位二进制数，取值范围为0~15，定义了文本右边阴影的厚度。当outline_style＝10时该字段才出现。

47. shadow_bottom

4位二进制数，取值范围为0~15，定义了文本底部阴影的厚度。当outline_style＝10时该字段才出现。

48. shadow_Y_component

5位二进制数，为文本阴影颜色定义了Y值，取值范围为0~31。当outline_style＝10时该字段才出现。

49. shadow_opaque_enable

1位标志位，值为1表示文本阴影颜色应为不透明的（无视频混合）。值为0表示将与视频做50%的混合。当outline_style＝10时该字段才出现。

50. shadow_Cr_component

5位二进制数，为文本阴影颜色定义了Cr值，取值范围为0~31。当outline_style＝10时该字段

才出现。

51. shadow_Cb_component

5位二进制数，为文本阴影颜色定义了Cb值，取值范围为0~31。当outline_style＝10时该字段才出现。

52. reserved_bits

3字节保留位，每个字节的值为0000 0000。当outline_style＝11时该字段才出现。 723

53. bitmap_length

16位二进制数，定义了其后压缩位图的字节数。

54. compressed_bitmap()

55. reserved_bits

[*n*]字节保留位，每个字节的值为0000 0000。只有当subtitle_type≠0001时该字段才出现。

56. descriptor_loop

[*n*]描述符，可能在descriptor_loop中出现。

57. CRC_32

32位CRC值。

13.23.2 DVB字幕

DVB字幕（ETSI EN 300 743）要比数字有线电视字幕复杂得多。

字幕流由stream_id＝私有流1的PES包传送。参考节目的节目时钟，它们的显示时间以显示时间标识（PTS）形式给出。一个字幕流传送一个或多个字幕服务。

每个字幕服务包含为特殊要求提供字幕所需的文本或图形。例如，可采用独立的字幕服务来传送不同语言类型的字幕。每个字幕服务在字幕页面序列中显示相关信息。

字幕页面覆盖在视频图像之上。一个字幕页面包含一个或多个字幕区。

每个字幕区是一个有位置、大小、像素深度和背景颜色等属性的矩形区域。字幕区作为背景框架能够放置一个或多个字幕对象。

一个字幕对象代表字符、词、文本行、整个句子、商标或图标。

PES数据字段语法如下。

1. data_identifier

8位字段，值为0x20，指示DVB字幕流。

2. subtitle_stream_ID

8位字段，值为0x00，指示DVB字幕流。

3. sync_byte

8位字段，值为0000 1111。

4. segment_type

8位字段，指示segment_data_field中包含的数据类型。为字幕定义下列segment_type值：

0x10＝页面构成段

0x11＝区域构成段

0x12＝CLUT定义段

0x13＝对象数据段

0x14＝显示定义段

0x80＝显示集结束段

5. page_ID

16位二进制数，标识了字幕服务。用于为一个字幕服务传送数据的段，其值与DVB字幕描述符中composition_page_ID匹配。用于传送可能由多字幕服务共享的数据的段，其值与ancillary_page_ID匹配。

724

一个频繁且经常优先采用的方法是，在不同的PID上采用不同的流传递截然不同的服务。

6. segment_length

16位二进制数，指示在segment_data_field中包含的字节数。

7. segment_data_field

这个字段是段的负载，系统定义了多个段类型。

页面构成段：承载关于页面构成的信息，如包括的区域列表、每一个区域的位置和页面的任何超时信息。

区域构成段：承载关于区域构成和属性的信息，如大小、背景颜色、像素深度、采用的颜色对照表（color lookup table、CLUT）以及在区域内这些组成物所包括的一个对象列表。

CLUT定义段：包含关于一个特定CLUT的信息，如用于一个CLUT条目的颜色。

对象数据段：承载关于特定文本或图形对象的信息。图形对象的对象数据段包含游程编码的位图颜色；对于文本对象，承载的是字符串编码。

显示集结束段：用来告知在当前显示集解码可以开始之前，不需要接收其他段了。

8. end_of_PES_data_field_marker

8位字段，值为1111 1111。

13.24 增强的电视节目安排

正如第8章中所讨论的，采用IP多播绑定的SMPTE 363M传输类型B广播数据是作为3部分来传输的，即声明、触发器和资源。声明在一个已知的多播IP地址和UDP端口上传输，并且指向可从特定多播IP地址和UDP端口上获得的触发器和资源。

725

SCTE 42规范定义了这些声明、触发器和资源是如何作为一个MPEG-2流的一部分而包括在其中的，如图13-26所示。在节目的PMT内标识这些数据（通过采用MAC地址列表描述符）实现数据与节目的关联。

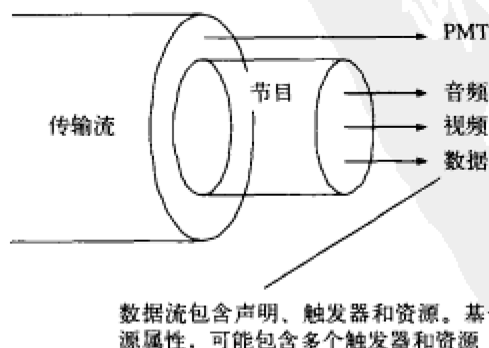


图13-26 承载在单个PID流上的声明、触发器和资源数据

在某些情况下，可能需要由有唯一PID的流来传送声明，而触发器和资源由另外一个有唯一PID的流来传送，如图13-27所示。这项技术的优势是，如果应用不能够接收触发器和资源，解码

器不需要处理与它们相关的IP数据报。声明总是要处理的，但是相比而言其数据量非常小。基于其各自特性，如语言类型、目标观众、资源大小和最大比特率等，触发器和资源也可能由不同的流传送。

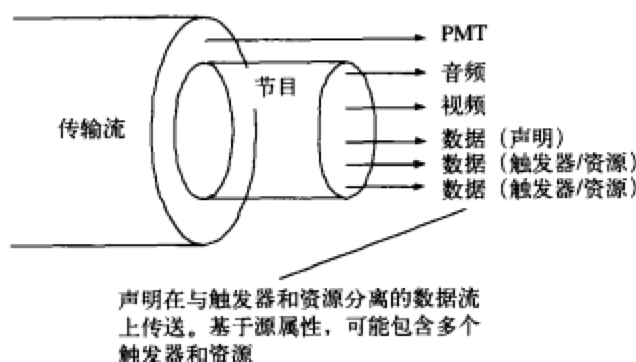


图13-27 声明与触发器和资源分开承载

图13-28给出了图13-27的一个扩展。触发器和资源流在多个节目之间共享。例如，两个节目可能希望从同一个天气服务获取数据。

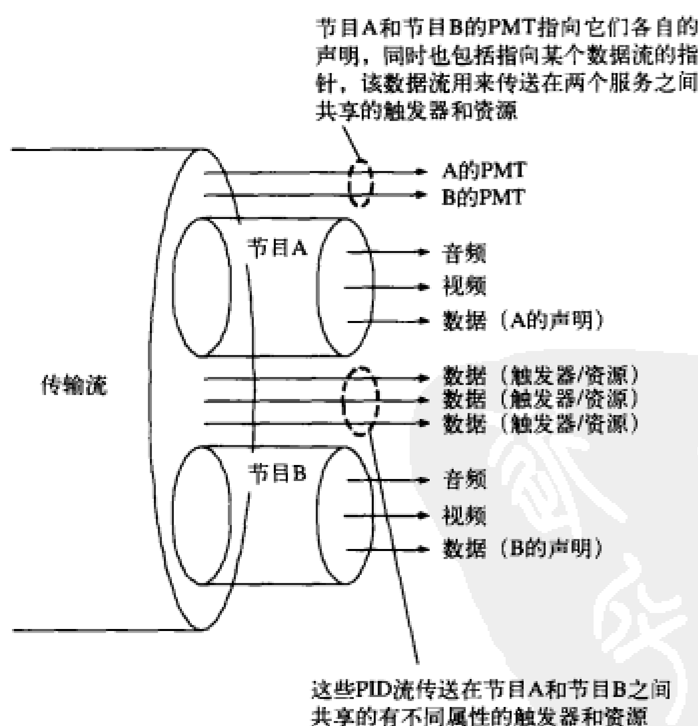


图13-28 共享触发器和资源

13.25 数据广播

MPEG-2通过DSM-CC规范（MPEG-2.6）支持各种内容分发工具和协议。可以利用DSM-CC工具的应用包括视频点播、数据广播、因特网接入和IP多播。

首先，DSM-CC简单地提供类似于VCR的功能（快进、快退、暂停等）作为MPEG-2.1的一个

附件。该功能后来扩展成MPEG-2.6，用来处理分布式内容的选择、访问和控制。因此，DSM-CC现在已构成了一个大的工具集：

- 网络会话和资源控制
- 客户端配置
- 客户端下载
- 流控制，文件访问
- 交互式和广播下载
- 数据和对象轮播
- 交换式数字广播信道变更协议

图13-29说明了ATSC和DVB数据广播如何利用DSM-CC实现各种数据广播特征。

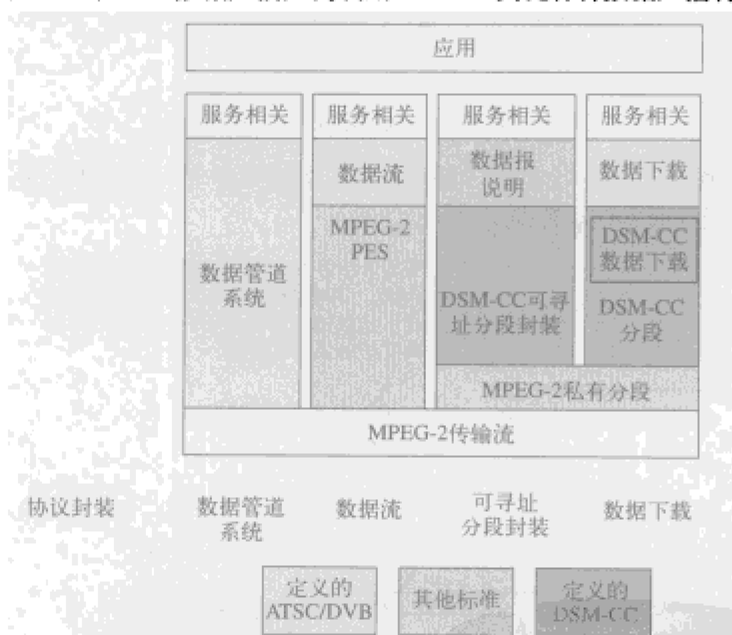


图13-29　封装纵览

13.25.1　轮播

轮播循环重复它们的内容。如果一个接收者希望从轮播上访问特别的数据，他只需要简单地等待下一次的数据广播。

数据轮播包含未定义的数据，因此，当这些数据被接收到之后，接收者需要知道怎样处理它们。数据轮播经常用于下载新的系统软件。

对象轮播包含可标识的数据（如数据流、图像、触发器事件、可执行的应用等）以及一个列出了轮播内所有对象的目录。对象轮播经常用于购物服务、电子节目指南（EPG）、广告和其他交互式功能。与数据轮播不同，对象轮播可以改变每个独立对象的循环速率。

726-727

13.25.2　在MPEG-2上的IP多播传输

IP多播基于DSM-CC在MPEG-2传输流上传递IP数据报。ATSC（A/92）和DVB（ETSI EN 301 192）采用形式稍微不同的LAN模拟来传递包数据。为数字有线电视系统设计的SCTE 42需要支持上述两种技术。

目前有两个主要的协议：DVB多协议封装（MPE）和ATSC DSM-CC可寻址分段。数字有线电视系统普遍采用其中一个或者别的协议，但是不同时采用这两个协议。然而，在一个节目内，封装改变可以在任何时间发生。

DVB实现是与DSM-CC分段包含私有数据兼容的。table_id＝0x3E，表示一个DSM-CC分段包含私有数据（参见表13-49），在这种情况下，该DSM-CC分段就是MPE数据报分段。

ASTC实现是同样与DSM-CC分段包含私有数据兼容的。table_id字段是0x3F，一个DSM-CC分段包含可寻址分段（参见表13-49）。

传送IP多播数据的每一个流有一个stream_type＝0x0D在PMT中与其相关联，表示它传送的是DSM-CC分段。

1. MAC地址的组成

IETF RFC 1112定义了一个IP多播地址到一个以太网MAC地址（ATSC称其为deviceId）的映射。通过将IP多播地址的低23位放置到MAC地址01：00：5E：xx：xx：xx（16进制）的低23位上，IP多播地址被映射到对应的硬件多播地址。IETF RFC 1700规定，MAC地址的位23总是为0。

2. 在MPEG-2上传输

图13-30说明了IP数据报是如何封装和分段为MPEG-2传输包的。IP数据报在IP层分割，因此它们不会超过定义的最大传输单元（Maximum Transfer Unit，MTU）大小，通常为4080字节。

一个单独的数据报分段可以横跨同一PID的多个MPEG-2包。同样，消息可以背靠背地放置在MPEG包载荷中。这需要使用pointer_field（PF）来指向下一个消息开始的位置。

MAC地址列表描述符用来（通过多播MAC组地址）标识由每一个流传送的数据。

13.25.3 数据广播机制

在一个MPEG-2传输流内，有大量的封装协议用来传输数据。前面已经讨论了IP多播。其他常见的技术包括：异步数据流化、同时数据流化和同步数据流化。

1. 异步数据流化

由DSM-CC分段（ATSC）或PES包（ARIB和DVB）传送的异步数据流，用于数据传输不受任何时序约束支配的应用。

对于ARIB和DVB的私有流2，PES包的stream_id＝0xBF。因为PES包的最大尺寸为64KB，为了满足传输流需求，PES包被分段成多个184字节的单元。

在PMT中，stream_type＝0x0D（ARIB和ATSC）或0x06（DVB）。

2. 同时数据流化

作为PES包载荷传送的同时数据流，用于需要连续流化数据以规则的和恒定的数据率到达接收者的应用。为了达到这个要求，时序信息被包含在流中。

因为PES包的最大尺寸为64KB，为了满足传输流需求，PES包被分段成多个184字节的单元。

对于私有流1，PES包的stream_id＝0xBD，允许使用PES头，包括显示时间戳（PTS）。然而，PTS的分辨率从11.1μs扩展到了74 ns。在PMT中，stream_type＝0x0D（ARIB）、0xC2（ATSC）或0x06（DVB）。

3. 同步数据流化

作为PES包载荷传送的同步数据流，用于需要精确数据显示、但并不必要显示规整次数的应用。显示次数通常与视频、音频或数据流相关联。

因为PES包的最大尺寸为64KB，为了满足传输流需求，PES包被分段成多个184字节的单元。

对于私有流1，PES包的stream_id＝0xBD，允许使用PES头，包括显示时间戳（PTS）。在

728～730

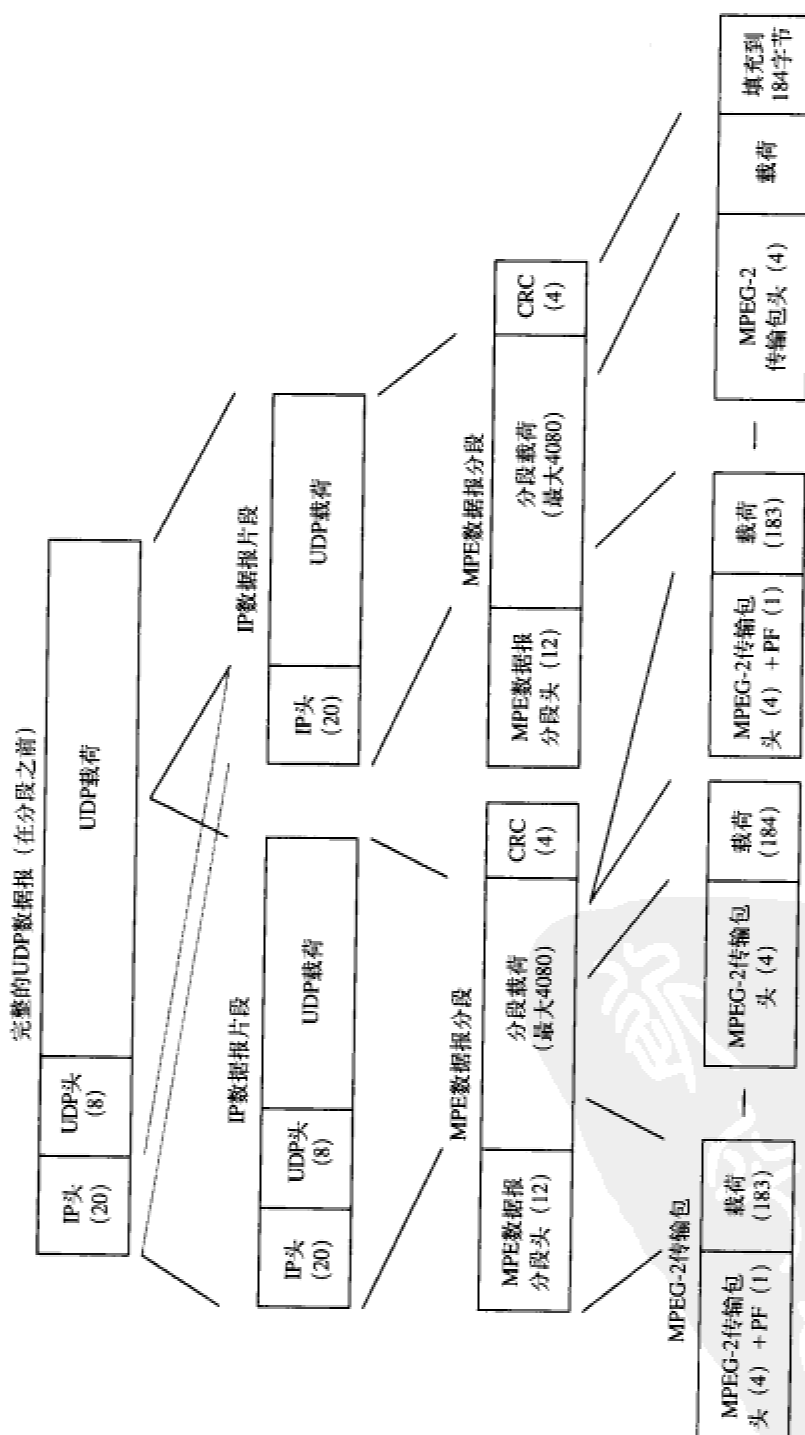


图13-30　在MPEG-2传输流上传送IP多播数据报

PMT中，stream_type＝0x0D（ARIB）或0x06（ATSC和DVB）。

4. 数据管道系统

数据管道系统是一个基本的异步传输机制，用于MPEG-2传输流上的数据（数据是直接被插入到MPEG-2传输流包的载荷中）。没有采用分区、表和PES结构。在这种机制中没有标准的分割和聚合数据报的方法，具体的实现方式由应用定义。

731

13.26 解码器注意事项

视频解码器本质上执行与编码器相反的功能。从编码的比特流中，解码器重构I帧。利用I帧、附加的编码数据和运动矢量，就可以生成P帧和B帧。最后，这些重构帧以合适的顺序输出。

图13-31举例说明了一个基本的MPEG-2视频解码器的框图。图13-32和图13-33分别说明了支持SNR和时域可分级的MPEG-2解码器的框图。

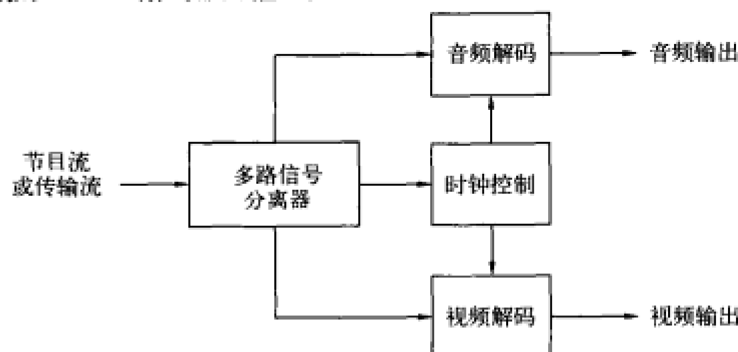


图13-31 简化的MPEG-2解码器框图

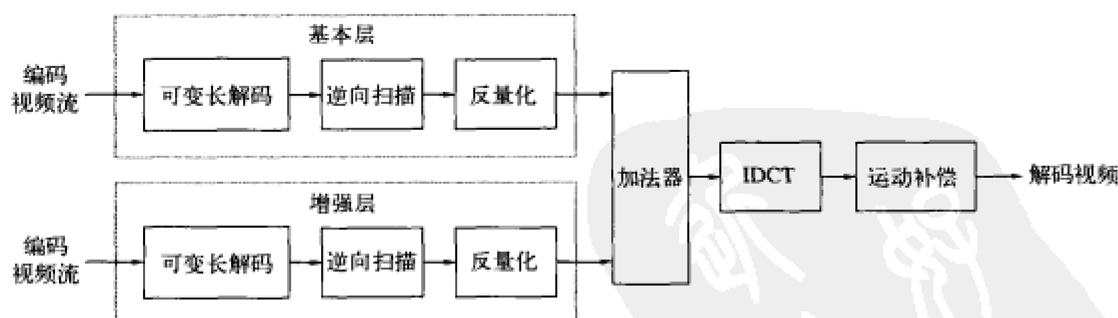


图13-32 简化的MPEG-2 SNR可分级解码器框图

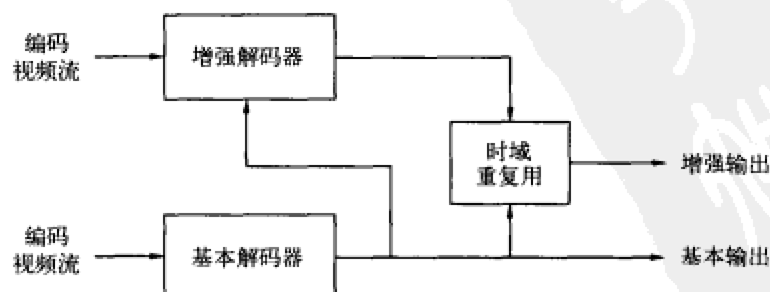


图13-33 简化的MPEG-2时域可分级解码器框图

音频和视频同步

MPEG-2解码器为每一个节目产生具有不同PID（包标识）的PES包。MPEG-2解码器仅识别具有选择节目PID的包，而忽略其他的包。

MPEG-2编码器包含了一个27MHz的振荡器和33位的计数器，称作系统时间时钟（system time clock，STC)。STC是一个由90kHz时钟所导出的33位值，该90kHz时钟是通过将27MHz时钟除以300而获得的。它属于一个特定的节目，并且是该节目视频和音频编码器的主时钟。

1. 时间戳

由于有可能出现双向预测编码，因此图像在压缩之后可能会乱序发送。每一个图像都有一个可变的数据量，且可能由于复用和传输而导致有可变的延时。为了保持音频和视频的同步，对于一幅图像，将周期性地发送时间戳。

MPEG-2编码器通过采样STC注意到一个输入图像或者音频块的发生时间（以及其编码输出的出现时间)。在此基础上，添加一个等于编码器和解码器缓冲区延时之和的常数值，从而得到一个33位的显示时间戳（PTS)。除此之外，还可能会添加一个33位的解码时间戳（DTS)，指示数据应该从MPEG-2解码器缓冲区取走缓冲和解码的时间。DTS和PTS是相同的，对双向预测（B）帧进行重排序的情况除外。

因为显示时间间隔是均匀的，因此并不总是需要包含一个时间戳（它们可以由解码器插入)，但是显示时间间隔不能超过700ms。唇同步是通过将时间戳同时合成到视频和音频PES包头中而获得的。

2. PTS和DTS

当B图像出现时，可能需要某个图像在该B图像出现之前先被解码，因此，它可以作为B图像的参考。尽管图像可以以IBBP的顺序显示，但是它们必须以IPBB的顺序传送。因此，这里存在两种类型的时间戳。

DTS指示一个图像必须解码的时间，而PTS指示的是该图像必须在MPEG-2解码器输出处出现的时间。B图像的解码和显示是同时的，因此它们仅包含PTS。当解码器接收到一个IPBB序列时，I图像和P图像必须在第一个B图像解码之前先解码。MPEG-2解码器一次只能解码一个图像，因此，I图像首先解码并存储起来，当解码P图像时，已解码的I图像将输出，从而其后可以跟随B图像。

PES包头中的PTS和DTS标志被设置来指示PTS单独出现或者PTS和DTS同时出现。音频PES包头仅包含一个PTS，因为音频PES包从不乱序传送，它们没有包含DTS字段。

比特流中包含的PTS（或DTS)，间隔不得超过700ms。ATSC进一步强制将PTS或DTS插入到每一个编码图像的起始处。

732

3. PCR

MPEG-2解码器的输出同样是带STC值的时间标识，称为PCR（节目时钟基准）或SCR（系统时钟基准)，两者用来同步MPEG-2解码器的STC与MPEG-2编码器的STC。在节目流中，时钟基准被称为PCR，而传输流中则被称作SCR。

传输流包头中的适配域周期性地用来包含PCR信息。MPEG-2要求每秒10PCR的最小传送数，DVB定义最小数是每秒25PCR。SCR要求出现的频率至少需达到每700ms一次。

❑ 同步

采用VCXO和PLL锁定MPEG-2解码器的27MHz时钟，并借此接收PCR，从而实现同步，如图13-34所示。这项技术保证在长时间连续工作期间，解码器的接收缓冲不会因源时钟稍快或稍慢于解码器时钟而过载或欠载。对于流视频应用，一般要求±100ppm的调整范围。

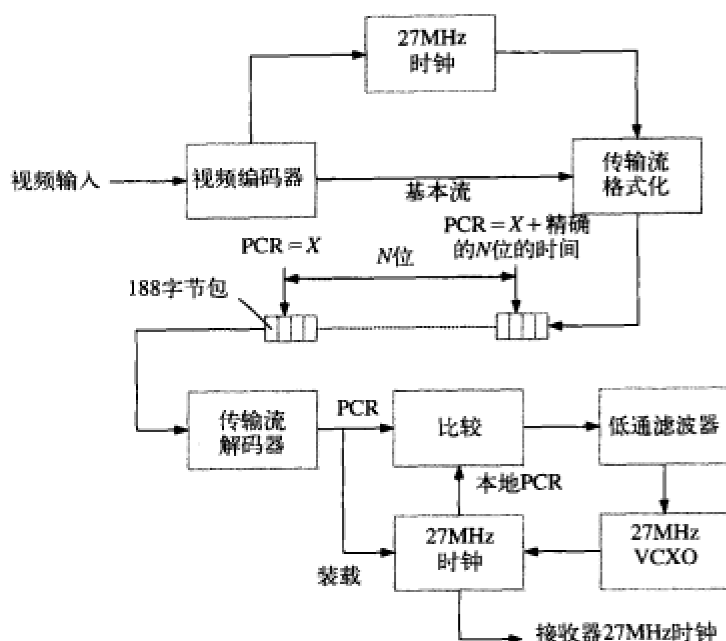


图13-34 由MPEG-2解码器生成的27MHz时钟

在MPEG-2解码器中，VCXO生成一个标称频率为27MHz的时钟，用来驱动本地PCR计数器。这个本地PCR与PES包头PCR比较，产生一个PCR相位误差。PCR相位误差过滤后控制VCXO，使得本地PCR计数器开始与PES包头PCR同步。非连续指示符可能用于复位本地PCR计数器，也可能被用来减少过滤，从而帮助MPEG-2解码器更快地锁定新的时序。

4. 唇同步问题

唇同步是与MPEG-2标准无关的实现问题。假设音频和视频在MPEG-2编码器输入中是同步的，PCR和PTS的支持提供了维护音频和视频时序关系所需的工具。而在目前已经开发出来的MPEG-2解码器中，这些工具并没有正确地实现，这种情况在MPEG-2编码器中也可能存在。

在MPEG-2解码器中没有正确地实现这些工具的一个代表性例子与PTS的读取有关。当首次接收到编码数据时，某些MPEG-2解码器只是简单地读取PTS，然后基于时域参考值不正确地快速向前移动，且不再检查PTS，直到通道发生变化或者重新接收数据为止。而其他解码器完全忽视时域参考值，取而代之的是关注时间戳。随着时间的推移，早期的实现错误逐步被纠正，从而减少唇同步问题的发生。

733～735

参考文献

1. *A Guide to MPEG Fundamentals and Protocol Analysis (Including DVB and ATSC)*, Tektronix, 1997.
2. ATSC A/53E, *ATSC Digital Television Standard*, September 2006.
3. ATSC A/54A, *Recommended Practice: Guide to the Use of the ATSC Digital Television Standard*, December 2003.
4. ATSC A/65C, *ATSC Standard: Program and System Information Protocol for Terrestrial Broadcast and Cable*, January 2006.
5. ATSC A/90, *ATSC Data Broadcast Standard*, July 2000.
6. ATSC A/91, *Recommended Practice: Implementation Guidelines for the ATSC Data Broadcast Standard*, June 2001.
7. ATSC A/92, *ATSC Standard: Delivery of IP Multicast Sessions over ATSC Data Broadcast*, January 2002.
8. CEA-708-B, *Digital Television (DTV) Closed Captioning*, December 1999.
9. Digital Video Magazine, "*Not All MPEGs Are Created Equal*," by John Toebes, Doug Walker, and Paul Kaiser, August 1995.

10. Digital Video Magazine, "*Squeeze the Most From MPEG,*" by Mark Magel, August 1995.
11. ETSI EN 300 468, *Digital Video Broadcasting (DVB): Specification for Service Information (SI) in DVB Systems*, May 2006.
12. ETSI EN 300 472, *Digital Video Broadcasting (DVB): Specification for Conveying ITU-R System B Teletext in DVB Bitstreams*, May 2003.
13. ETSI EN 300 743, *Digital Video Broadcasting (DVB): Subtitling Systems*, November 2006.
14. ETSI EN 301 192, *Digital Video Broadcasting (DVB): DVB Specification for Data Broadcasting*, November 2004.
15. ETSI EN 301 775, *Digital Video Broadcasting (DVB): Specification for the Carriage of Vertical Blanking Information (VBI) Data in DVB Bitstreams*, May 2003.
16. ETSI TS 101 154, *Digital Video Broadcasting (DVB): Implementation guidelines for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*, June 2005.
17. ISO/IEC 13818-1, *Generic coding of moving pictures and associated audio information, Part 1: Systems.*
18. ISO/IEC 13818-2, *Generic coding of moving pictures and associated audio information, Part 2: Video.*
19. ISO/IEC 13818-3, *Generic coding of moving pictures and associated audio information, Part 3: Audio.*
20. *OpenCable™ Host Device 2.0 Core Functional Requirements*, July 2005.
21. SCTE 20 2004, *Methods for Carriage of Closed Captions and Non-Real Time Sampled Video.*
22. SCTE 21 2001R2006, *Standard for Carriage of NTSC VBI Data in Cable Digital Transport Streams.*
23. SCTE 27 2003, *Subtitling Methods for Broadcast Cable.*
24. SCTE 42 2002, *IP Multicast for Digital MPEG Networks.*
25. Watkinson, John, *The Engineer's Guide to Compression*, Snell and Wilcox Handbook Series.

736~737



第14章 MPEG-4和H.264

MPEG-4是在MPEG-2的成功和经验基础之上建立起来的，它最为人们所熟知的特点有：

- 比MEPG-2低的比特率（对于同样的视频质量）；
- 使用自然的或合成的对象，且这两种对象可以一起渲染从而生成场景；
- 支持交互行为。

对创作人员而言，MPEG-4将使内容的制作具有更好的重用性、更大的灵活性，并受到更有利的保护。

对消费者而言，MPEG-4能提供更多的交互行为，并且由于其具有比MPEG-2更低的比特率，因此，使得用户能够享受新的网络（例如DSL）和移动产品上的内容。

MPEG-4是一项ISO标准（ISO/IEC 14496），目前由以下19个部分组成。

系统	ISO/IEC 14496-1
视觉	ISO/IEC 14496-2
音频	ISO/IEC 14496-3
一致性测试	ISO/IEC 14496-4
参考软件	ISO/IEC 14496-5
DMIF	ISO/IEC 14496-6
参考软件	ISO/IEC 14496-7
IP网络传输	ISO/IEC 14496-8
参考硬件	ISO/IEC 14496-9
高级视频（H.264）	ISO/IEC 14496-10
场景描述	ISO/IEC 14496-11
ISO文件格式	ISO/IEC 14496-12
IPMP扩展	ISO/IEC 14496-13
MP4文件格式	ISO/IEC 14496-14
H.264文件格式	ISO/IEC 14496-15
动画扩展	ISO/IEC 14496-16
文本流格式	ISO/IEC 14496-17
字体压缩	ISO/IEC 14496-18
合成纹理流	ISO/IEC 14496-19

MPEG-4通过使用描述元（descriptive element）提供了一个标准的方式来表示音频、视频甚至是图像媒体对象（例如，采用描述元取代图像实际比特来描述图像）。媒体对象可以是自然的，也可以是合成的（计算机生成），并且其表现可以独立于周围环境或背景。

MPEG-4同样描述了如何融合多个媒体对象来生成一个场景。与传统的发送图像比特不同，MPEG-4发送的是媒体对象并且接收者能据此组成图像。这样一个机制使得该标准具有以下特点：

- 对象可以被放置在任何地方；
- 可以对对象进行几何变化；
- 可以对对象进行分组；
- 修改数据的属性以及转换数据；

738

- 动态地改变一个场景的视角。

14.1　音频概述

MPEG-4音频支持各种不同的应用，从简单的语音到多通道高质量的音频。

音频对象（音频编解码器）采用特定的工具（tools）组合来有效地表示不同类型的音频对象。*档次*（profile）采用特定的音频对象类型组合来有效地为特定市场需求服务。*等级*（levels）在一个档次里详细定义了大小、速率以及复杂度限制等来确保互操作性。

目前，大多数的解决方法仅仅支持少数几个最流行的音频编解码器（通常是AAC-LC和HE-AAC），而不是一个或多个档次/等级。

14.1.1　一般音频对象类型

一般音频对象类型支持大范围的质量、比特率以及声道数。对于自然音频，MPEG-4支持AAC编码（Advanced Audio Coding）、BSAC编码（Bit Sliced Arithmetic Coding）以及TwinVQ（Transform Domain Weighted Interleave Vector Quantization）算法。可用的音频对象如下所示。

1. AAC-Main对象

AAC-Main对象在MPEG-2 AAC-Main的基础上增加了感知噪声修整（Perceptual Noise Shaping，PNS）工具。

2. AAC-LC对象

AAC-LC（Low Complexity）对象在MPEG-2 AAC-LC的基础上增加了PNS工具。AAC-LC也有错误恢复（Error Resilient）版本，即ERAAC_LC。

3. AAC_SSR对象

AAC_SSR（Scalable Sampling Rate）对象在MPEG-2 AAC_SSR的基础上增加了PNS工具。

4. AAC-LTP对象

AAC-LTP（Long Term Predictor）对象和AAC-LC对象相似，它采用长期预测器（Long Term Predictor，LTP）来取代AAC-LC预测器。在确保相同效率的情况下，明显降低了实现代价。AAC-LTP同样有错误恢复版本，即ER AAC-LTP。

5. AAC-Scalable对象

AAC-Scalable对象允许大量的可伸缩组合。它们仅仅支持单声道或双声道立体声。AAC-Scalable同样有错误恢复版本，即ER AAC-Scalable。

6. ER AAC-LD对象

ER AAC-LD（Low Delay）派生于AAC，并且支持所有双声道或更多的声音声道。它们支持48kHz以上的采样率，并且使用512或480采样点的帧长（相对于AAC采用的1024或960采样点）来获得20ms的最大算术时延。

7. ER BSAC对象

ER BSAC对象取代了无噪声编码（AAC量化光谱数据）和比例因子。ER BSAC采用一个基本层比特流和许多小的增强层比特流来确保实时调整服务质量。

739

8. HE-AAC对象

HE-AAC（High Efficiency）对象是AAC和频带复制SBR（Spectral Band Replication）技术的结合，被设计用于极低比特率的编码，例如32kbit/s的立体声。

9. TwinVQ对象

TwinVQ对象基于固定速率的矢量量化，从而取代了AAC中的Huffman编码。它们针对的是比

AAC还低的比特率，支持单声道和立体声。TwinVQ同样有错误恢复版本（ER TwinVQ）。

14.1.2 语音对象类型

语音编码工作的比特率范围在2~24kbit/s之间。当采用可变速率编码时，可能会出现更低的比特率，例如平均1.2kbit/s。已有的音频对象如下。

1. CELP对象

当比特率在4~24kbit/s之间时，CELP（Code Excited Linear Prediction）对象支持8kHz和16kHz的采样率。CELP对象同样有错误恢复版本（ER CELP）。

2. HVXC对象

当固定的比特率为2~4kbit/s（低于2kbit/s使用不同的比特率模式）时，HVXC（Harmonic Vector eXcitation Coding）对象支持8kHz的单声道语音，它具有在解码中改变倾斜度和速度的能力。HVXC对象也有错误恢复版本（ER HVXC）。

14.1.3 合成语音对象类型

可伸缩的语音合成（Text-to-Speech，TTS）对象提供了一个低比特率（200~1200bit/s）的语音音位表示。叙述型的内容能简单地生成而不用记录自然语音。TTS接口允许语音信息以国际音标字母（International Phonetic Alphabet，IPA）或者以任何语言的文本形式（书面）传输。合成语音同样可以和面部动作对象同步。

14.1.4 合成音频对象类型

合成音频支持是由结构化音频解码器（Structured Audio Decoder）实现所提供的，它使得基于乐谱的应用程序可以控制信息发送到用特殊语言描述的乐器。可用的音频对象如下。

1. 主要合成对象

主合成对象（Main Synthetic Object）可以使用所有的MPEG-4结构化音频工具。它们支持SAOL（Structured Audio Orchestra Language）音乐合成语言和采用SASBF（Structured Audio Sample-Bank Format）的波表合成。

2. 波表合成对象

波表合成对象是主合成对象的子集，采用了SASBF和乐器数字接口（Musical Instrument Digital Interface，MIDI）工具。它们提供了相对简单的采样合成。

3. 一般MIDI对象

一般MIDI对象提供了与现有内容的互操作性。

740

14.2 视觉概述

MPEG-4视觉被分为两个部分。其中，一部分是在此要讨论的MPEG-4.2，它包含了早期MPEG-4视频编解码器。另一部分是MPEG-4.10，它详细定义了“先进视频编解码器”，即通常所说的H.264，这部分内容将在本章末尾进行讨论。

MPEG-4的视觉规范在三个比特率范围内进行了优化，分别是：低于64kbit/s、64~384kbit/s、0.384~4Mbit/s。

对于高质量的应用，通过采用与低比特率情况下相同的工具和比特流语法，获得更高的比特率也是可能的。

在MPEG-4中，视觉对象（视频编解码器）采用特定的工具（tool）组合来有效地表示不同类

型的视觉对象。*档次*（profile）采用特定的视觉对象类型组合来有效地为特定市场部分服务。*等级*（Levels）在档次中指定了大小、速率以及复杂度限制等来确保互操作性。

目前，由于硅的成本问题，大部分的解决方法仅仅支持少数几个MPEG-4.2视频编解码器。随着MPEG-4.10（H.264）和SMPTE421M（VC-1）的引入，人们对MPEG-4.2视频编解码器的兴趣将急剧减少。这两个方案将使MPEG-4.2的性能翻一番。

14.2.1　YCbCr颜色空间

4：2：0的YCbCr颜色空间被用于大多数的*视觉对象*。每个分量可以用4~12位来表示，最常见的是采用8位。

为了支持在编辑过程中所需的更高图像质量，MPEG-4.2的Simple Studio和Core Studio对象可以采用4：2：2、4：4：4、4：2：2：4以及4：4：4：4：4的YCbCr或者RGB采样选项。

和H.263以及MPEG-2一样，MPEG-4.2视频编解码器同样是基于宏块、块和DCT变化的。

14.2.2　视觉对象

MPEG-4使用自然的和合成的*视觉对象*，从而取代了早期MPEG规范中采用的视频帧（frame）或者图像（picture）。在给定时间点的视频对象实例被称作*视觉对象平面*（visual object plane，VOP）。

与MPEG-2非常类似，MPEG-4也有I（内部的）、P（预测的）和B（双向预测的）VOP。S-VOP是sprite对象的VOP。S（GMC）-VOP采用基于过去某个参考VOP全局运动补偿的预测编码。

在MPEG-4中可以采用任意形状的视频对象，也可以采用矩形对象。例如，一个MPEG-2视频流可以是一个矩形视频对象。

视觉对象同样可以是可伸缩的，从而使得可以从整个比特流中的几块来重构有用的视频。这是通过采用一个基本层和一个或多个增强层来实现的。

由于当前*自然视觉对象类型*已经成为市场上最受关注的话题，因此这里只讨论自然视觉对象类型。

14.2.3　MPEG-4.2自然视觉对象类型

MPEG-4.2支持多个自然视觉对象类型（视频编解码器），表14-1列出了其中几个令人感兴趣的对象类型。更普遍的对象类型如下。

741

表14-1　常用MPEG-4.2自然视觉对象类型可用工具

工　具	对象类型						
	主对象	核心对象	简单对象	高级简单对象	高级实时简单对象	高级编码效率对象	细粒度可伸缩对象
VOP类型	I、P、B	I、P、B	I、P	I、P、B	I、P	I、P、B	I、P、B
色度格式	4：2：0	4：2：0	4：2：0	4：2：0	4：2：0	4：2：0	4：2：0
隔行扫描	×	—	—	×	—	×	×
全局运动补偿（GMC）	—	—	—	×	—	×	—
四分之一像素运动补偿（QPEL）	—	—	—	×	—	×	—
条带重同步	×	×	×	×	×	×	×

（续）

工　具	对象类型						
	主对象	核心对象	简单对象	高级简单对象	高级实时简单对象	高级编码效率对象	细粒度可伸缩对象
数据分区	×	×	×	×	×	×	×
可逆VLC	×	×	×	×	×	×	×
短头	×	×	×	×	×	×	×
方法1和方法2量化	×	×	–	×	–	×	×
形状自适应DCT	–	–	–	–	–	×	–
动态分辨率转换	–	–	–	–	×	×	–
NEWPRED	–	–	–	–	×	×	–
二元形状	×	×	–	–	–	×	–
灰度形状	×	–	–	–	–	×	–
Sprite	×	–	–	–	–	–	–
细粒度可伸缩性（FGS）	–	–	–	–	–	–	×
FGS时域可伸缩性	–	–	–	–	–	–	×

742

1. 主对象

主对象提供最高的视频质量。与中心对象相比，主对象同样支持灰度形状（grayscale shape）、sprite以及交织与累进的内容。

2. 核心对象

核心对象采用的是主对象所使用工具的一个子集，尽管仍然支持B-VOP。它们通过发送额外的P-VOP，也同样支持可伸缩性。二元形状（binary shape）可以包含一个固定的透明度，但不能实现由灰度形状编码所提供的可变透明度。

3. 简单对象

简单对象为低比特率、有错误恢复能力、任意高宽比（aspect ratio）的矩形自然视频对象。简单对象使用的工具是中心对象使用工具的一个子集。

4. 高级简单对象

高级简单对象在仅仅支持矩形对象时看起来和简单对象是很相似的，但是通过增加以下少数工具可以提高其效率：B帧、1/4像素运动补偿（QPEL）和全局运动补偿（GMC）。

5. 细粒度可伸缩对象

细粒度可伸缩对象能使用最多8个可伸缩层，因此，传送质量可以很容易地适应传输和解码环境。

14.2.4 MPEG-4.2自然视觉档次

MPEG-4.2支持许多视觉档次和等级。由于自然视觉档次是目前市场中最受关注的，因此这里仅讨论自然视觉档次（表14-2和表14-3）。更普遍的档次如下。

1. 主要档次

主要档次（Main Profile）是用于广播应用的，同时支持交织与累进的内容。它把最高质量的视频和任意形状的对象连接起来了。

2. 核心档次

核心档次（Core Profile）对更高质量的交互服务是非常有用的，它在保证高质量的同时确保

了有限的复杂度，并且支持任意形状的对象。这个档次同样可以支持移动广播服务。

3. 简单档次

简单档次（Simple Profile）专门用于低复杂度应用。主要应用是移动服务和Internet。

4. 高级简单档次

高级简单档次（Advanced Simple Profile）提供了在大比特率范围下分发基于帧的单层视频的能力。

5. 细粒度可伸缩档次

743

细粒度可伸缩档次（Fine Granularity Scalable Profile）专门用于Internet流媒体和无线多媒体。

表14-2　MPEG-4.2自然视觉档次和级别。4:4:4:4:4:4意味着4:4:4RGB+3个辅助通道；4:2:2:4YCbCr表示4:2:2YCbCr+1个辅助通道

MPEG-4.2档次	支持的形状	备　注	级别	典型分辨率	最大对象数目	最大比特率
主要档次	任意的	附加的工具和功能	L4	BT.709	32	38.4Mbit/s
			L3	BT.601	32	15Mbit/s
			L2	CIF	16	2Mbit/s
核心档次	任意的	附加的工具和功能	L2	CIF	16	2Mbit/s
			L1	QCIF	4	384kbit/s
高级核心档次	任意的	较高的编码效率	L2	CIF	16	2Mbit/s
			L1	QCIF	4	384kbit/s
N比特	任意的		L2	CIF	16	2Mbit/s
简单档次	矩形		L3	CIF	4	384kbit/s
			L2	CIF	4	128kbit/s
			L1	QCIF	4	64kbit/s
高级简单档次	矩形	较高的编码效率	L5	BT.601	4	8Mbit/s
			L4	352 × 576	4	3Mbit/s
			L3b	CIF	4	1.5Mbit/s
			L3	CIF	4	768kbit/s
			L2	CIF	4	384kbit/s
			L1	QCIF	4	128kbit/s
			L0	QCIF	1	128kbit/s
高级实时简单档次	矩形	较高的差错恢复能力	L4	CIF	16	2Mbit/s
			L3	CIF	4	384kbit/s
			L2	CIF	4	128kbit/s
			L1	QCIF	4	64kbit/s
核心可伸缩档次	任意的	空域和时域可分级性	L3	BT.601	16	4Mbit/s
			L2	CIF	8	1.5Mbit/s
			L1	CIF	4	768kbit/s
简单可伸缩档次	矩形	空域和时域可分级性	L2	CIF	4	256kbit/s
			L1	CIF	4	128kbit/s
			L0	QCIF	1	128kbit/s
细粒度可伸缩档次	矩形	SNR和时域可分级性	L5	BT.601	4	8Mbit/s
			L4	352 × 576	4	3Mbit/s
			L3	CIF	4	768kbit/s
			L2	CIF	4	384kbit/s
			L1	QCIF	4	128kbit/s
			L0	QCIF	1	128kbit/s

（续）

MPEG-4.2档次	支持的形状	备 注	级别	典型分辨率	最大对象数目	最大比特率
高级编码效率档次	任意的	较高的编码效率	L4	BT.709	32	38.4Mbit/s
			L3	BT.601	32	15Mbit/s
			L2	CIF	16	2Mbit/s
			L1	CIF	4	384kbit/s
核心专用档次（采用10位像素数据）	任意的	附加的工具和功能	L4	BT709，60P，4：4：4 BT.709，30I，4：4：4：4：4：4	16	900 Mbit/s
			L3	BT.709，30I，4：4：4 BT.601，4：2：2：4	8	450Mbit/s
			L2	BT.709，30I，4：2：2 BT.601，4：4：4：4：4：4	4	300 Mbit/s
			L1	BT.601，4：2：2：4 BT.601，4：4：4	4	90Mbit/s
简单专用档次（采用10位或者12位像素数据）	任意的		L4	BT709，60P，4：4：4 BT.709，30I，4：4：4：4：4：4	1	1800 Mbit/s
			L3	BT.709，30I，4：4：4 BT.709，30I，4：2：2：4	1	900Mbit/s
			L2	BT.709，30I，4：2：2 BT.601，4：4：4：4：4：4	1	600Mbit/s
			L1	BT.601，4：2：4 BT.601，4：4：4	1	180Mbit/s

表14-3 通用MPEG-4.2档次支持的对象

MPEG-4.2对象类型	MPEG-4.2档次						
	主要档次	核心档次	简单档次	高级简单档次	高级实时简单档次	高级编码效率档次	细粒度可伸缩档次
主对象	×	–	–	–	–	–	–
核心对象	×	×	–	–	–	×	–
N比特	–	–	–	–	–	–	–
简单对象	×	×	×	×	×	×	×
高级简单对象	–	–	–	×	–	–	×
高级实时简单对象	–	–	–	–	×	–	–
高级编码效率对象	–	–	–	–	–	×	–
细粒度可伸缩对象	–	–	–	–	–	–	×

744～745

14.3 图形概述

图形档次详细说明了BIFS工具中哪个图形元件能被用来建立场景。尽管图形在系统规范中已经定义了，但实际上它只不过是另一种媒体档次，就像音频和视频一样，所以在这里也将对它进行详细的讨论。

MPEG-4定义了4个层次的图形档次：简单2D、完整2D、完整和3D音频图形。它们的区别在于由解码器所支持的BIFS工具中的图形元件不同，如表14-4所示。

简单2D档次提供在场景中放置一个或多个视觉对象所需的基本特征。

完整2D档次提供2D图形函数并支持类似于任意的2D图形和文本的特征，它有可能和视觉对

象关联在一起。

表14-4　MPEG-4图形档次支持的图形元件（BIFS工具）

BIFS工具的图形元件	图形档次			图形工具（BIFS节点）	图形档次		
	简单2D	完整2D	完整		简单2D	完整2D	完整
显示	×	×	×	雾	–	–	×
方框	–	–	×	字形	–	×	×
位图	×	×	×	带索引的面谱集合	–	–	×
背景	–	–	×	带索引的2D面谱集合	–	×	×
2D背景	–	×	×	带索引的线集合	–	–	×
圆	–	×	×	带索引的2D线集合	–	×	×
色彩	–	×	×	线特性	–	×	×
圆锥体	–	–	×	材料	–	–	×
坐标	–	–	×	2D材料	–	×	×
2D坐标	–	×	×	普通工具	–	–	×
2D曲线	–	×	×	像素文本	–	×	×
圆柱体	–	–	×	点光源	–	–	×
有方向的光线	–	–	×	点集合	–	–	×
高程网格	–	–	×	2D点集合	–	×	×
表达式	–	–	×	矩形	–	×	×
挤压	–	–	×		×	×	×
面谱	–	–	×	球体	–	–	×
延期面谱网络	–	–	×	点光	–	–	×
延期面谱表	–	–	×	文本	–	×	×
延期面谱转换	–	–	×	文本坐标	–	×	×
FAP	–	–	×	文本转换	–	×	×
FDP	–	–	×	视位	–	–	×
FIT	–	–	×				

完整档次提供诸如高程网格（elevation grid）、挤压（extrusion）和复杂光照等高级功能。它使得能够用完全的虚拟世界来展现高度的现实主义。

3D音频图形档次能够用来定义场景的听觉特性（几何学、听觉吸收、扩散和物质透明度）。该档次在对音频信号进行环境均衡的应用上十分有用。

14.4　视觉层

一个MPEG-4视觉场景由一个或多个视频对象所组成。目前，最普遍的视频对象是简单的矩形视频帧。

每一个视频对象可能有一个或多个层来支持时域或空域可伸缩编码。这样就能使视频能够以层次化的方式来重构，即以一个基本层开始并在上增加一定数量的增强层。其中，当需要更高的可伸缩性时，例如当一副图像被映射成一个2D或3D对象时，可以采用小波变换来进行。

视觉比特流提供了对场景的层次化描述。层次的每一个等级都可以通过比特流中的唯一起始码来访问。见图14-1。

14.4.1　视觉对象序列

视觉对象序列（VS）是一个完整的场景，它包含所有的2D或3D、自然的或合成的对象以及任何增强层。

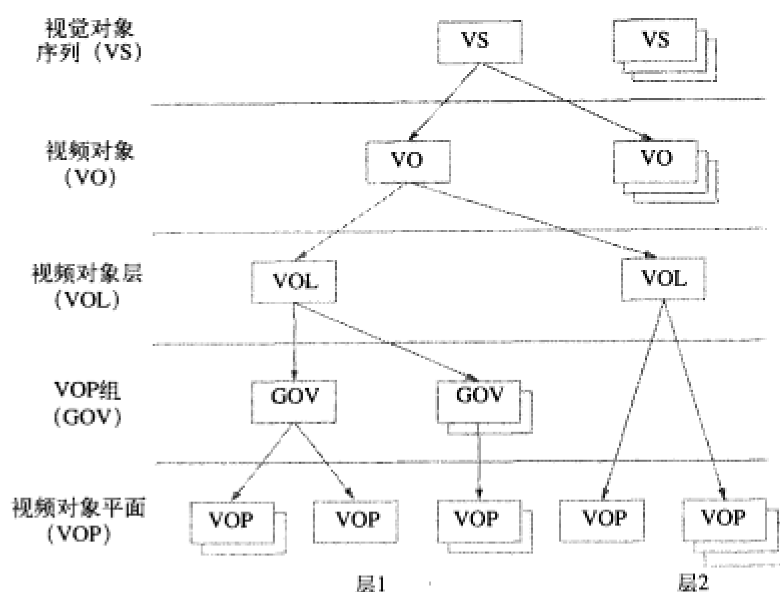


图14-1 MPEG-4视频比特流逻辑结构示例

14.4.2 视频对象

一个视频对象（VO）对应场景中的一个特定对象。在大多数简单情况下，VO可以是一个矩形帧，或者它可以是与场景中某一对象或背景对应的任意形状的对象。

14.4.3 视频对象层

每一个视频对象可以采用可伸缩方式（多层）或非伸缩方式（单层）来编码，并且采用视频对象层（VOL）来表示，而具体采用何种方式取决于特定的应用所需。VOL支持可伸缩编码。一个视频对象能采用空域或时域上的可伸缩性（从低到高的分辨率）来编码。取决于可用带宽、计算能力以及用户偏好等参数，上述所需的分辨率解码器都能使用。

746～747

MPEG-4有两种类型的视频对象层：提供全部MPEG-4功能性的视频对象层和提供精简功能且头更短的视频对象层。后者提供了与基准H.263兼容的比特流。

14.4.4 视频对象平面组

748

每一个视频对象都是在时域采样的，视频对象的每一次时域采样都是一个视频对象平面。视频对象平面可以组合在一起来形成视频对象平面组（GOV）。

GOV将视频对象平面组合在一起，它能提供比特流中视频对象平面相互独立编码的点，从而提供随机访问比特流的点。GOV是可选的。

14.4.5 视频对象平面

视频对象平面（VOP）是视频对象的一次时域采样。VOP能相互独立地编码，或者通过运动补偿相互依赖地编码。传统的视频帧能用矩形的VOP来表示。

14.5 对象描述框架

和MPEG-2不同的是，MPEG-4并不把多个基本流复合形成一个单独的传输流或节目流。

每个对象（音频、某个视觉对象的某层等）的数据、场景描述信息（用来声明对象的时空联系）以及对象控制信息在单独的基本流里传输。合成对象可以通过使用BIFS来生成，以提供图形和音频。BIFS比场景描述语言含有更多的信息，它把自然的和合成的对象结合在同一合成空间。

*对象描述框架*是用来识别、描述并将基本流相互联系起来的*对象描述符*集，它还用于描述在场景中使用的对象，如图14-2所示。

一个初始的对象描述符，派生于对象描述符，包含两个描述符。一个描述符指向*场景描述（基本）流*，另一个指向对应的*对象描述符（基本）流*。

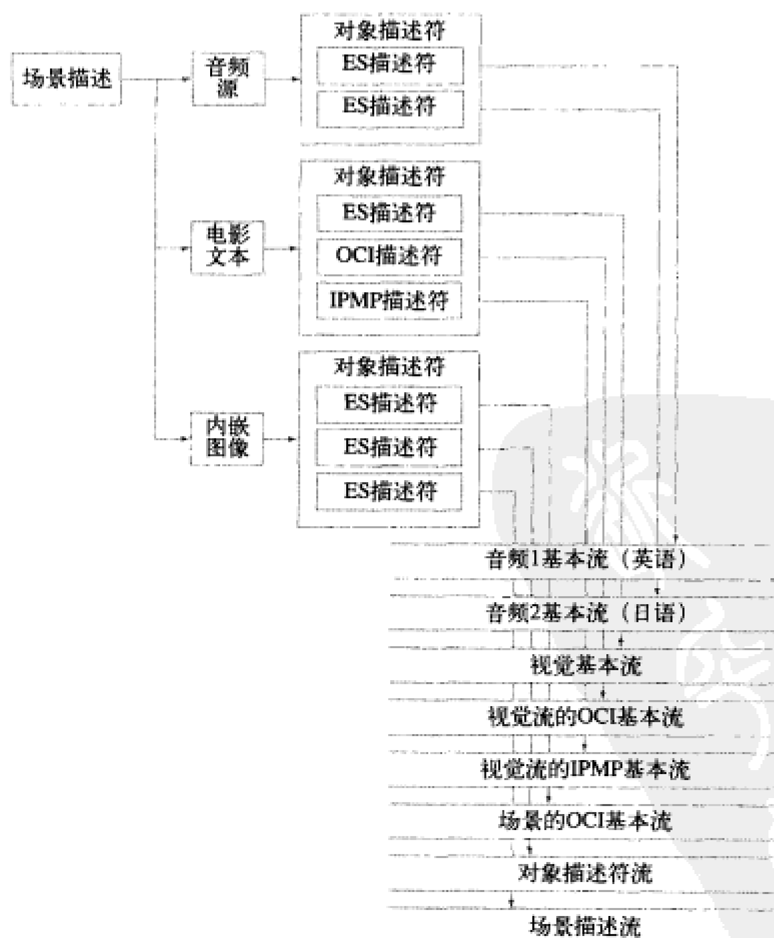
14.5.1　对象描述符流

对象描述符（OD）由一个专门的基本流（术语称为对象描述符流）传输。

对象描述符能有效地将相关基本流集合联合在一起，所以它们被解码器视为一个单独的实体。每一个对象描述符包含其他描述符，通常这些其他描述符指向一个或多个与单个节点和单个音频或视觉对象相关联基本流。因此，这便支持了多个可选的流，比如各种不同的语言。

另外，一个对象描述符还能指向辅助的数据，例如*对象内容信息*（OCI）和*知识产权管理与保护*（IPMP）。

749



750

图14-2　将基本流链接到场景

对象描述符不仅仅是简单地一个接一个存在于对象描述符流里。更准确地说，它们是被封装

在对象描述符命令（object descriptor command）里的。这些命令使得对象描述符能被动态地传送、更新或者删除无效流引用。这样一来当某个对象的一个新的基本流可用时，它能被及时告知，或者当某个基本流不再可用时，能够将参考移动到其中。更新是以时间戳的形式来指示它们何时起作用。时间戳和其他基本流被放置在同步层。

14.5.2 对象内容信息

对象内容信息（OCI）基本流用于传送OCI事件。每一个OCI事件都是由OCI描述符组成的。

OCI描述符用于描述关联对象的一些特征，例如关键词、内容的文本描述、语言、父母级[①]、创建日期、作者等。

如果OCI信息一成不变，很可能会用对象描述符流里的CCI描述符来传送这些信息。

14.5.3 知识产权管理与保护

IPMP基本流用于向一个或多个IPMP系统传送IPMP信息。IPMP系统提供在接收端知识产权管理和内容保护功能。

如果IPMP信息很少发生改变，很可能会用对象描述符流里的IPMP描述符来传送这些信息。

14.6 场景描述

为了在接收端组合一个多媒体场景，简单地发送多个数据流是不够的。例如，对象可能会被放置在2D或3D空间里，每一个对象有自身的坐标系统。当物体被放置在一个场景中时，会将它们自身的坐标系统转换为场景的坐标系统。因此，在接收端为用户组合成实际有意义的场景就需要额外的数据。这些额外的数据被称为*场景描述*。

场景图形元件（BIFS工具）用于描述视听的原始数据和属性。图14-3举例说明了这些形成层次场景图形的场景图形元件以及它们之间的关系。场景图形不一定是静态的，而且根据需要元件可以增加、删除或者修改。

场景图形类定义了允许的可能被采用的场景图形元件集合。

BIFS

二进制格式场景（Binary Format For Scenes，BIFS）不仅用于描述场景构成信息，也用于描述图形元件。BIFS和VRML之间的本质区别是BIFS采用*二进制格式*，而VRML采用的是*文本格式*。BIFS支持由VRML定义的元件以及某些VRML不支持的元件，包含压缩二进制格式、流、流化动画、2D原语、增强音频以及面部动作。

751

1. 压缩二进制格式

BIFS支持对场景图形信息的有效二进制表示。编码可以是有损的或无损的。有损压缩是基于上下文知识的：如果某些场景图形数据已经接收到了，就可以预见接下来可能收到的图形数据的类型和格式。

2. 流

BIFS是设计用于场景的传输，传输的形式是初始场景加上场景中的某些修改。

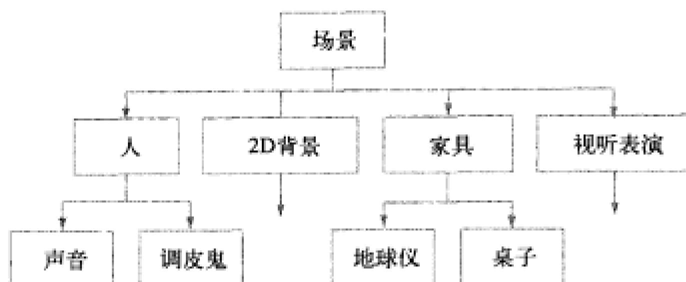
3. 流化动画

BIFS包含用于连续动画改变场景中元件数值的低开销（low-overhead）方法。这为BIFS和VRML支持的分类机提供了另一选择。

① 美国影视分级之一。——译者注

4. 2D原语

BIFS本质上支持2D场景采用低复杂度、低消耗方案，如传统的电视。与将世界分为2D和3D不同，BIFS能同时将2D和3D元件融合在一个场景里。



752

图14-3　MPEG-4层次场景图示例

5. 增强音频

BIFS通过使用音频场景图形改进了对音频的支持，使得音频源被混合或者生成声音效果。

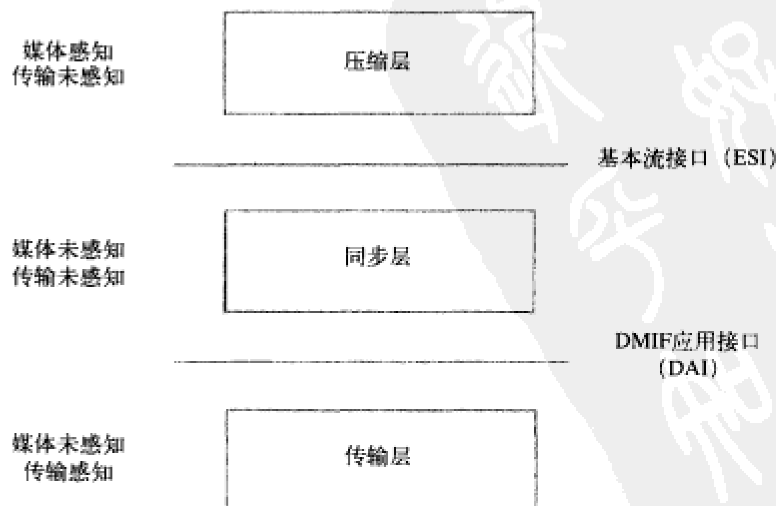
6. 面部动作

BIFS将动态面部特性提升到场景等级。这使得其成为场景中的一员，可以与其他的BIFS特性有效地集成，类似于其他的视听对象。

14.7　基本流同步

14.7.1　同步层

同步层（图14-4）将每个基本流打包成访问单元的序列，访问单元是同步层中的唯一需要端到端保护的语义实体，它们的内容是不透明的。访问单元将作为同步的基本单位，它们通过使用灵活的语法能对相关特性进行编码，产生SL数据包。然后，这些SL数据包被传递到传输层。由基本流产生的SL数据包序列被称为RT打包流（SPS，SL-packetized stream）。



753

图14-4　MPEG-4压缩层、同步层和传输层之间的关系

和MPEG-2 PES不同的是，同步层不是一个独立的流。相反，它是一个被映射到特定传输层的中间格式，例如IP、MPEG-2传输流等。由于这个原因，在同步层里就没有必要包含起始码、SL数据包长度等信息，这些信息都已经被包含在传输层协议里了。

SL数据包用于两个目的。第一，访问单元在适应特定传输层时能以任何方式打包。第二，它使得编码器在已经了解传输层特性（如最大传输单元MTU大小）的情况下，能够有效地指导打包过程。

多路基本流的同步是由承载对象时钟参考（OCR）、解码时间戳（DTS）、组合时间戳（CTS）以及在同步层里的时钟参考等实现的。

同步层语法是灵活的，它可以为每个基本流单独配置。例如，低比特流音频流可能需要最小报头的时间戳，高比特率的音频流可能需要十分准确的时间戳。

14.7.2 DMIF应用接口

与MPEG-2不同的是，MPEG-4支持多路并发使用方案（如本地检索、远程交互、广播、多播等）以及多路并发传输技术。

多媒体传输集成框架（Delivery Multimedia Integration Framework，DMIF）应用接口或者DAI是在传输和接收过程中控制数据在同步层和传输层（图14-4）之间交换的接口。它能够访问、表现和同步采用不同技术（如MPEG-2传输流、IP多播、RTP等，甚至是同时采用这些技术）发送或接收的MPEG-4内容。

因此，控制和配置数据到特定传输或存储协议（又可被称为*载荷格式规范*）的规范必须和管理传输层规范的组织联合实现。例如，在IP上传输MPEG-4的开发工作就是和IETF联合完成的。

14.8 基本流的多路复用

关于MPEG-4内容的传输通常没有在规范之中给出定义。

然而，对现有传输协议的分析表明，需要一个附加的复用层来复用多路基本流。偶然突发和低比特率的MPEG-4流有时需要映射到采用固定长度数据包（例如MPEG-2传输流）或者高多路复用开销（例如RTP/UDP/IP）的传输层。然而，大量的传输流可能带来的潜在问题是，在管理和代价方面存在较重的负担。

为解决这一问题，MPEG-4定义了一个简单复用包语法——FlexMux。它允许复用多个SL打包流到一个独立的低开销FlexMux流。

另外，已经实现了将SL数据包封装到普通传输层协议的规范，其中包括MPEG-2传输和节目流、IP（参见第19章）以及MPEG-4文件格式等。 754

14.8.1 FlexMux

FlexMux将一个或多个瞬时比特率变化的SL打包流复用到*FlexMux数据包*，FlexMux数据包是FlexMux的基本数据实体，其数据包长度是可变的。

FlexMux通过FlexMux信道数来识别来自不同基本流的SL数据包。每个SL打包流都被映射到一个FlexMux信道。因此，包含不同SL打包流数据的FlexMux数据包可以任意交叉。安排在一个流中的FlexMux数据包序列称为一个FlexMux流。

14.8.2 在MPEG-2之上的MPEG-4

MPEG-2 PES是封装内容的通用命名。MPEG-4定义了对SL数据包以及在PES包里的FlexMux

流的封装。

一个SL打包流被映射到一个PID或者MPEG-2复用的stream_id。每个SL数据包被映射到一个PES包。仅仅传输PES报头里的信息和从SL数据包报头中去除重复数据，可以减少PES和SL数据包报头冗余。

FlexMux数据包的整数序号同样也可以在PES包中传输，从而进一步减少复用开销。多个SL打包流可以被映射到一个MPEG-2 PID或者stream_id。自从可以采用SL数据包报头中传输的时间戳信息进行同步之后，PES报头几乎不再使用。

14.8.3 MP4文件格式

为了交换MPEG-4内容，定义了新的文件格式。这种文件格式支持元数据，以方便索引、快速搜索和随机访问。

14.9 知识产权管理与保护（IPMP）

IPMP，又可被称作数字版权管理（DRM），为实现知识产权权利管理提供了一个接口和工具，而不是一个完整的系统。

管理和保护的等级和类型的提供是由内容的价值和商业模式来决定的。由于这个原因，对IPMP系统的完整的设计留给了应用开发者。

这一架构使得可以同时采用公开和私有方案，同时允许互操作，支持采用多于一种类型的保护（即加密、水印、版权管理等），并支持使用已定义的设备交互信息在设备之间进行内容转移（如在家庭网络中进行内容分发）。

对于受保护的内容，在表现之前IPMP工具需求将告知解码器。工具配置和初始化信息由IPMP描述符或者IPMP基本流所传输。所需求的工具能被嵌入、下载或者通过其他方式来获得。

755

在IPMP描述符里的控制点和排序序列信息允许不同的工具在系统的不同地方发生作用。IPMP数据被IPMP描述符或者IPMP基本流所传输，其中还包含版权容器、关键字容器和工具初始化数据。

14.10 MPEG-4.10（H.264）视频

ITU-T H.264即为先前众所周知的H.26L、JVT、JVT codec、AVC和先进视频编码。现在，H.264是两个最新视频编码器中的一个，另一个是基于微软Windows Media Video 9编码器的SMPTE 421M（VC-1）。H.264已经作为第10部分而被并入MPEG-4规范中了。

H.264采用了许多新的工具用于改善性能。这些新的工具包括：支持8比特、10比特和12比特4：2：2 和4：4：4 YCbCr，整数变换，UVLC、CAVLC和CABAC熵编码，多重参考帧，内部预测，内循环解块过滤器，SP和SI条带，许多新的错误恢复工具。

14.10.1 档次和级别

和其他的视频编码相似的是，档次详细指定了语法（即算法），级别详细指定了不同的参数（解决方案、帧率、比特率等）。不同的级别如表14-5所示。

1. 基本档次（BP）

基本档次设计用于逐行扫描视频，如视频会议、基于IP的视频以及移动应用等。基本档次所用的工具包括：I和P条带类型、1/4像素运动补偿、UVLC和CAVLC熵编码、任意条带顺序（ASO）、灵活的宏块顺序（FMO）、冗余条带（RS）、4：2：0 YcbCr格式。

需要注意的是，基本档次不仅仅是主要档次的一个子集。许多解决方案实现了基本档次不带ASO或者FMO的一个子集，也是主要档次（更容易实现）的一个子集。

2. 扩展档次（XP）

扩展档次主要用于移动和因特网流媒体应用。除基本档次的工具之外，还包括如下工具：B、SP和SI条带类型，条带数据分割，加权预测。

756

表14-5 MPEG-4.10（H.264）级别。MB＝宏块，MV＝运动矢量

级别	最大MB/s	最大帧大小（MB）	典型帧分辨率	典型帧率（帧/秒）	每两个连续MB之间的最大MV	最大参考帧数	最大比特率
1	1 485	99	176×144	15	–	4	64 kbit/s
1.1	3 000	396	176×144 320×240 352×288	30 10 7.5	–	9 3 3	192 kbit/s
1.2	6 000	396	352×288	15	–	6	384 kbit/s
1.3	11 880	396	352×288	30	–	6	768 kbit/s
2	11 800	396	352×288	30	–	6	2 Mbit/s
2.1	19 800	792	352×480 352×576	30 25	–	6	4 Mbit/s
2.2	20 250	1 620	720×480 720×576	15 12.5	–	5	4 Mbit/s
3	40 500	1 620	720×480 720×576	30 25	32	5	10 Mbit/s
3.1	108 000	3 600	1280×720	30	16	5	14 Mbit/s
3.2	216 000	5 120	1280×720	60	16	4	20 Mbit/s
4	245 760	8 192	1920×1080 1280×720	30 60	16	4	20 Mbit/s
4.1	245 760	8 192	1920×1080 1280×720	30 60	16	4	50 Mbit/s
4.2	491 520	8 192	1920×1080	60	16	4	50 Mbit/s
5	589 824	22 080	2048×1024	72	16	5	135 Mbit/s
5.1	983 040	36 864	2048×1024 4096×2048	120 30	16	5	240 Mbit/s

757

3. 主要档次（MP）

主要档次主要用于大范围内的广播应用。在基本档次之上的另外的工具包含：隔行编码，B条带类型，CABAC熵编码，加权预测，4：2：2和4：4：4 YCbCr（10比特和12比特格式），不支持ASO、FMO和RS。

4. 高级档次（HP）

在初始的规范完成之后，增加了逼真度范围扩展（FRExt）改善。从而向规范里加入了以下四种附加的档次。

- 高级档次（HP）：增加对在4×4和8×8块大小之间自适应选择的支持，其目的是为选择亮度空间转换和用于转换系数的编码器特定、频率相关的缩放比例矩阵。
- 高级10档次（Hi10P）：增加对9比特或10比特4：2：0 YCbCr的支持。
- 高级4：2：2档次（Hi422P）：增加对4：2：2 YCbCr的支持。

- 高级4：4：4档次（Hi444p）：增加对11比特或12比特采样，4：4：4 YCbCr或RGB，残余颜色变换和可预测无损编码等的支持。

14.10.2　补充的增强信息

补充的增强信息（SEI）在与解码、显示或其他目的相关的过程中提供辅助。SEI包括：缓冲周期、图像时序、饼状扫描矩形、过滤器有效载荷、用户数据注册、用户数据注销、恢复点、解码参考图像标记副本、备用图像、场景信息、子序列信息、子序列层特征、子序列特征、全图像冻结、全图像冻结释放、渐进精细片段开始、渐进精细片段结束和运动受限条带组集。 758

逼真度范围扩展（FRExt）改善增加了三个新的补充增强信息（SEI）：影片颗粒特征、去块过滤器显示偏好和立体视频。

14.10.3　视频编码层

1. YCbCr颜色空间

H.264使用YCbCr颜色空间，支持4：2：2、4：2：2和4：4：4采样。4：2：2和4：4：4采样选项在4：2：0采样之上提升了色度分辨率，从而提高了图像质量。另外对于8比特YCbCr数据，H.264支持10比特和12比特YCbCr数据来进一步提高图像质量。

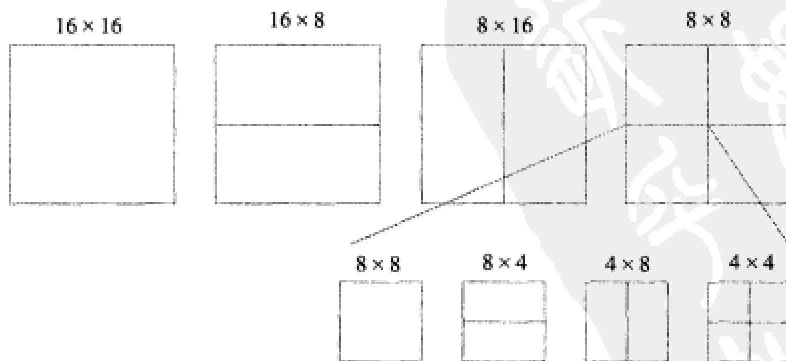
图3-8~图3-10中给出了H.264的4：2：0的采样结构。图3-2和图3-3中给出了4：2：2和4：4：4的采样结构。

2. 宏块

在H.264中，扩展了对16×16宏块的分割，如图14-5所示。

如此精细的尺度使得一个宏块能有大数量的运动矢量（至少32个），并对一定数量的块进行插值（至少96个）。为了限制编码器和解码器的复杂度，对连续的两个宏块所含的运动矢量的数量有所限制。

FMO提高了对错误的隐藏能力，它将宏块分配到另一个条带，从而使得它们能在非扫描序列中传输。这样便减少了一个错误将影响一大片空间区域的可能性，并通过使用邻域宏块预测丢失的宏块来改善错误的隐藏能力。



759 图14-5　为进行运动补偿而进行的H.264宏块分割。上：宏块分割。下：8×8部分分割

3. 内循环去块过滤器

H.264增加了内循环去块过滤器。它去除了由于相邻宏块有不同估计类型和不同的量化尺度所导致的块效应。过滤器同样也去除了由于相邻宏块有不同的变换或量化和运动矢量所导致的块效应。

循环过滤器使用内容自适应非线性过滤器来修改在块某侧或宏块边界的两个采样点。

4. 条带

自从它称为基本的独立空间元素起，条带在H.264里扮演着更加重要的角色。它阻止了在一个条带中的错误影响到另一个条带。这种灵活性允许扩展I图像、P图像和B图像类型降低到条带等级，从而产生了I、P和B条带类型。

任意条带顺序（ASO）允许条带被无序传输和接收。这改善了视频会议和网络应用中的低延迟现象。

残余条带同样能用于错误恢复。这种可选择的数据能被用来恢复任意损坏的宏块。

❑ SP条带和SI条带

除了I条带、P条带和B条带之外，H.264还增加了对SP条带（交换P）和SI条带（交换I）的支持。SP条带采用运动补偿预测，利用时间冗余来对条带进行重建（即使当采用不同参考条带时，也是如此）。SI条带利用了空间预测来同样重建一个对应的SP条带。

使用S条带能进行有效的比特流转换、随机访问、快进以及错误恢复，如图14-6和图14-7所示。

5. 内部预测

当运动估计无效时，可使用内部预测消除空间冗余。这项技术尝试基于相邻块来预测当前块，然后，对预测块和实际块之间的差异进行编码。这个工具对于通常存在许多空间冗余的单调背景十分有用。

6. 运动补偿

❑ 1/4像素运动补偿

大多数早期视频编码使用的1/2像素精度可用来改善运动补偿精度。H.264同样支持用于最新的MPEG-4视频编码中的1/4像素精度。

❑ 多重参考帧

H.264增加了对多重参考帧的支持。因此，改善预测过程增加压缩率，并使用丢失事件里另外的参考帧来提高错误恢复能力。

一个简单的宏块能使用至多8个参考帧（对HDTV至多3个），一个帧里的参考帧数量限制为16个。

为了补偿在目前和参考帧间的不同时间距离，预测块采用加权参数来作平均。这些参数不仅可以嵌入在流里，还可以是解码器从时间参考里隐式地推导获得的。 760

❑ 非受限的运动搜索

这项技术允许采用图像外的参考帧。丢失的数据可以利用边界数据通过空域预测获得。

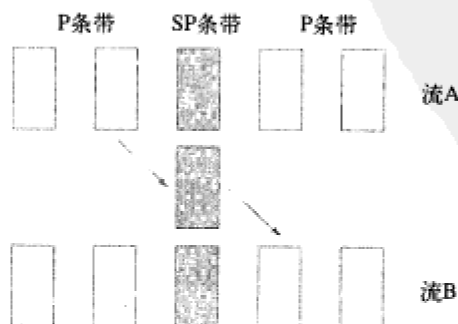
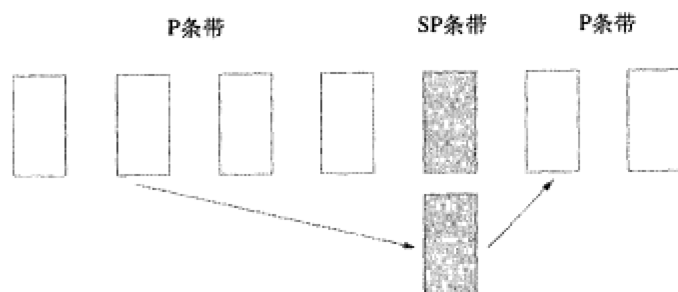


图14-6 采用SP条带切换到另一个流



761

图14-7 采用SP条带进行快进

7. 变换、缩放和量化

H.264采用简单的4×4整数变换。相反，早期的视频编码采用8×8 DCT，该变换操作在浮点系数上。附加的2×2变换被应用到4个CbCr DC系数。内部的16×16宏块有一个额外的4×4变换，这个4×4变换是为16个Y DC系数执行的。

因为H.264采用了更小的块，所以块状和环状效应得以减少。另一方面，使用整数系数消除了舍入误差（rounding error）。这些舍入误差将产生漂移效应，而漂移效应在基于DCT视频编码器中是普遍存在的。

对于量化，H.264采用52个均匀标量量化器集，其中步长依次递增12.5%。

量化之后，通过采用一个或两个扫描命令将从低频到高频依次扫描量化系数。

8. 熵编码

在量化和Z形扫描之后，H.264采用两种类型的熵编码：可变长编码（VLC）和内容自适应二元算术编码（CABAC）。

除了变换系数之外，H.264采用单个通用VLC（ULVC）表，该表使用无限扩展码字集（Exponential Golomb）。与其他视频编码器采用多重VLC表不同的是，在H.264中仅映射到单个UVLC表是根据统计学而定制的。

对于消耗大量带宽的变换系数，H.264采用自适应内容可变长度编码（CAVLC）。基于先前处理的数据，可以选择最好的VLC表。

通过采用自适应内容二元算术编码（CABAC），可额外提高5%~10%的效率。CABAC不断地更新新来数据的统计，并通过一个称为内容建模的过程实时自适应调整算法。

14.10.4 网络抽象层

网络抽象层（NAL）便于将H.264数据映射到不同的传输层，包含针对有线和无线的网络服务的RTP/IP、文件格式（如MP4）、服务于视频会议的H.32X和MPEG-2系统。

数据被组织到NAL单元，数据包包含整数数量字节。每个NAL的首字节表示的是有效载荷数据类型和剩余字节包含有效载荷数据。有效载荷数据可能与其他的数据交叉存取以阻止意外地产生起始码前缀。

当采用数据分区时，每一个条带被分成3个独立的部分，每一部分采用一个特定的NAL单元类型。这使得数据分区能作为一种有效的分层方法，用于将数据分离到不同重要性的等级。通过将数据分区到不同的NAL单元里，可以更容易地实现对数据不同部分采用不同的差错保护方法。

762

参考文献

1. *H.264 Advanced Video Coding: A Whirlwind Tour*, by PixelTools, 2003.
2. *H.264 Coding Efficiency Has a Price,* by Eric Barrau, Philippe Durieuz, and Stephane Muta, Sophia Antipolis Micro-Electronics Forum, January 2003.
3. ISO/IEC 14996-1, *Information Technology—Coding of Audio—Visual Objects—Part 1: Systems.*
4. ISO/IEC 14996-2, *Information Technology—Coding of Audio—Visual Objects—Part 2: Visual.*
5. ISO/IEC 14996-3, *Information Technology—Coding of Audio—Visual Objects—Part 3: Audio.*
6. ITU-T H.264, *Advanced Video Coding for Generic Audiovisual Services*, May 2003.
7. *Technical Overview of H.264/AVC*, by R. Schafer, T. Wiegand and H. Schwarz, EBU Technical Review, January 2003.
8. *The H.264/AVC Video Coding Standard for the Next Generation Multimedia Communication*, by M. Mahdi Ghandi and Mohammad Ghanbari, IAEEE invited paper.

763



第15章　ATSC数字电视

ATSC（Advanced Television System Committee）数字电视（DTV）广播标准应用于美国、加拿大、韩国、墨西哥、阿根廷这些国家。

其他三个主要DTV标准是DVB（Digital Video Broadcast）、ISDB（Integrated Services Digital Broadcasting）和OpenCable™。这些标准传输音/视频的基本能力类似，不同之处在于所采用的RF调制策略以及非音/视频服务的分辨率等级。表15-1对不同的ATSC标准进行了比较。

ATSC标准实际上是由一组标准组成：

A/52 数字音频压缩（AC-3和EAC-3）标准
A/53 ATSC数字电视标准
A/57 ASC传输的内容识别和标记
A/64 数字电视的传输测量和规范
A/65 地面广播和线缆的节目与系统信息协议
A/70 地面广播的条件接收系统
A/76 节目级的元数据通信协议
A/80 卫星传输数字电视（DTV）节目的调制和编码要求
A/81 直接到户的卫星广播标准
A/90 数据广播标准
A/92 采用数据广播标准的IP组播会议传输
A/93 同步和异步触发标准
A/94 数据应用参考模式
A/95 传输流文件系统标准
A/96 交互信道协议
A/97 软件下载数据服务
A/100 DTV应用软件环境：级别1（DASE-1）
A/101 先进通用应用平台（ACAP）
A/110 分布式传输同步标准

ATSC标准采用MPEG-2传输流在6MHz的单信道上传输压缩数字视频、压缩数字音频和数据。MPEG-2传输流可以包含多个视频流、多个音频流和/或数据。例如，可以同时包含一个节目的高清和标清数据流，并包含数据信息，如本地天气预报。

MPEG-2传输流的最大比特率为19.4Mbit/s（6MHz，空中信道）或38.8Mbit/s（6MHz，数字线缆信道）。

在所提供节目的数量和视频质量、分辨率之间进行权衡，19.4Mbit/s的比特率可以有多种灵活的应用方式。例如：

（1）HDTV节目
（1）HDTV节目＋（1）SDTV节目＋数据
（4）SDTV节目

除了在标准中增加E-VSB（在后续章节中讨论）以支持更健壮的操作模式之外，A-VSB的相关工作也正在开展中。

为了更好地适应移动通信市场并且和DVB-H和DMB相抗衡，A-VSB将会提高动态多路跟踪技术，以实现分层（结构化）调制、支持时分复用和帧分片技术。为了提高地面覆盖范围，A-VSB将会使得单频率网络（SFN）内不同信号塔之间的广播信号时序同步更加方便。

表15-1 ATSC标准之间的比较

参数	ATSC-T（地面）	ATSC-C（线缆）	ATSC-S（卫星）	ATSC-T E-VSB（地面）
视频压缩	MPEG-2			MPEG-2 MPEG-4.10 (H.264)
音频压缩	Dolby®数字音频			Dolby®数字音频 Dolby®数字音频增强版
复用	MPEG-2 传输流			
调制	8-VSB	16-VSB[1]	QPSK, 8PSK	采用ATSC-T
信道带宽	6MHz	6MHz	–	–

765

注：大多数数字线缆系统采用QAM而不是16-VSB。

15.1 视频能力

虽然，只要不超过最大比特率，视频节目可以采用任意分辨率，但通常所采用的分辨率只有表15-2所列的几种。大多数分辨率均允许采用隔行和逐行两种图像格式。

视频压缩是基于MPEG-2的。然而，对于某些MPEG-2参数有更细化的要求，这一点将在后续的章节中讨论。支持采用MPEG-4.10（H.264）直至HP@L4.0的标准正在制定中。

表15-2 ATSC视频常见有效分辨率

有效分辨率（Y）	SDTV或HDTV	帧速（p＝逐行，i＝隔行）			
		23.976p 24p	29.97i 30i	29.97p 30p	59.94p 60p
480×480	SDTV	×	×	×	×
528×480		×	×	×	×
544×480		×	×	×	×
640×480		×	×	×	×
704/720×480		×	×	×	×
1280×720	HDTV	×		×	×
960×1080		×	×	×	
1280×1080		×	×	×	
1440×1080		×	×	×	
1920×1080		×	×	×	

15.2 音频能力

音频压缩采用Dolby®数字音频并且支持1~5.1声道。

主音频，或可以构成完整服务（包含所有必要的节目单元）的关联音频，具有的比特率小于等于 448 kbit/s（典型情况下采用的比特率为384 kbit/s）。单声道包含单个节目单元的关联服务的比特率小于等于128 kbit/s。双声道仅包含对话的关联服务的比特率小于等于192 kbit/s。一个主音频和关联音频的组合具有的比特率不超过576 kbit/s。组合中的主音频和关联音频可以同时解码。

定义的音频服务有以下几种。

15.2.1　主音频服务：完整主音频（CM）

这种音频服务包含一个完整的音频节目（对话、音乐和声效）。音频所提供的服务通常就是这种服务，并且可以包含1~5.1声道。

采用VI、HI、C或VO关联服务可以进一步增强CM服务。并且，通过提供多CM服务，每个采用一种不同的语言，就可以提供多语种音频服务。

15.2.2　主音频服务：音乐和音效（ME）

这种音频服务包含一个音频节目的音乐和音效，但不包含对话。它可以包含1~5.1声道。基本的节目对话（如果有的话）通过D服务提供。

15.2.3　关联服务：针对视力障碍（VI）人群

这种服务通常包含对节目内容的叙述性描述。VI服务采用单个音频通道。VI和CM同时解码，使得视障人群也可以享受节目。

除了以一个单独的、包含叙述性描述的声道提供VI服务外，还可以在一个完整的主音频中加入音乐、音效、对话以及叙述性描述。这时，整个音频服务可能会用到5.1声道。

15.2.4　关联服务：针对听力障碍（HI）人群

这一服务通常只包含用于和CM服务一起重构的对话。这时，HI作为一个单独的声道。

除了以一个单独的声道提供HI服务外，还可以在一个完整的主音频中加入音乐、音效以及更易理解的对话。这时，整个音频服务可能会用到5.1声道。 766

15.2.5　关联服务：对话（D）

这种服务包含和ME服务联合使用的节目对话。

同时解码D和ME服务，并且将D服务混入ME服务的主声道，才能形成一个完整的音频节目。

如果ME服务包含两个以上的音频通道，D服务将是单声道的。如果ME服务包含两个声道，D服务也可以包含两个声道。这时，同时解码D和ME服务、将D和ME服务的左声道混合并将D和ME服务的右声道混合可以形成一个完整的音频节目。这种音频节目是包含音乐、音效和对话的双声道立体声信号。

通过提供多个D服务（每种语言一个）以及单个ME服务可以形成多语言音频节目。这种方法的效率要比通过提供多个CM服务高得多。然而，当ME服务中包含两个以上的音频通道时，对话只能包含在中央声道中。

15.2.6　关联服务：解说（C）

解说服务类似于D服务，不同之处在于，D服务传输的是必要的节目对话，而解说服务传递的是采用单音频通道的可选的节目解说。

此外，还可以在一个完整的节目中加入音乐、音效、对话以及解说。这时，整个音频服务可能会用到5.1声道。

15.2.7　关联服务：画外音（VO）

767 这种服务是一个单声道的服务，用于和ME服务混合并在ME解码过程中解码。

15.3 节目和系统信息协议

采用MPEG-2传输流之后可以获得足够的带宽以支持多个低带宽、非电视的服务。这些服务包括节目导航、字幕、天气预报、股市指数、热点新闻、软件下载、按次付费信息等。外加非电视服务（虚拟信道）的数量可以很容易地到达十个或者更多。服务的数量和类型将不断变化。

开发节目和系统信息协议（PSIP）的目的是为了采用更灵活、兼容性更强的方式支持这些非电视服务。PSIP是一个层次化关联表的小集合（见图15-1和表15-3），其目的是扩展MPEG-2 PSI表。它描述了在一个特定的MPEG-2传输流中传输的所有虚拟信道的信息。此外，模拟广播信道的信息也可以包含在内。

表15-3 ATSC PSIP表、描述符和描述符位置

描述符	描述符标识	地面广播表									
		PMT	MGT	VCT	RRT	EIT	ETT	STT	DCCT	DCCSCT	CAT
PID		每PAT	0x1FFB	0x1FFB	0x1FFB	每MGT	每MGT	0x1FFB	0x1FFB	0x1FFB	0x0001
table_ID		0x02	0xC7	0xC8	0xCA	0xCB	0xCC	0xCD	0xD3	0xD4	0x80,0x81(ECM) 0x82~0x8F(EMM)
重复率		400ms	150ms	400ms	1min	0.5s	1min	1s	400ms	1h	
AC-3音频流	1000 0001	M				M					
ATSC CA	1000 1000			O		O					
ATSC私有信息*	1010 1101										
CA	0000 1001	M									M
字幕服务	1000 0110	M				M					
组件名称	1010 0011	M									
内容分级	1000 0111	M				M					
内容标志符	1011 0110	O				M					
DCC到达请求	1010 1001								M		
DCC离开请求	1010 1000								M		
增强信令	1011 0010	M PMT-E									
扩展通道名称	1010 0000			M							
流派	1010 1011					M					
再发行控制	1010 1010	M				M					
服务位置	1010 0001			M							
SRM参考	0000 1001										M
填充*	1000 000										
时移服务	1010 0010			M							

注：M＝当存在时，该表中需要，O＝在该表中也可能存在，*＝没有要求。

15.3.1 所需表

1. 事件信息表（EIT）

事件信息表总共有128个，包括EIT0~EIT127，每个EIT描述在VCT中列出的每个虚拟通道的事件或电视节目。每个EIT的有效时间是3小时。由于总共有128个EIT，未来16天内的节目都可以提前预告。前4个EIT是必需的（建议的个数为24）。

EIT提供的信息包括开始时间、持续时间、标题、指向事件可选描述文本的指针、分级

(advisory) 数据、字幕服务数据、音频服务描述符等。

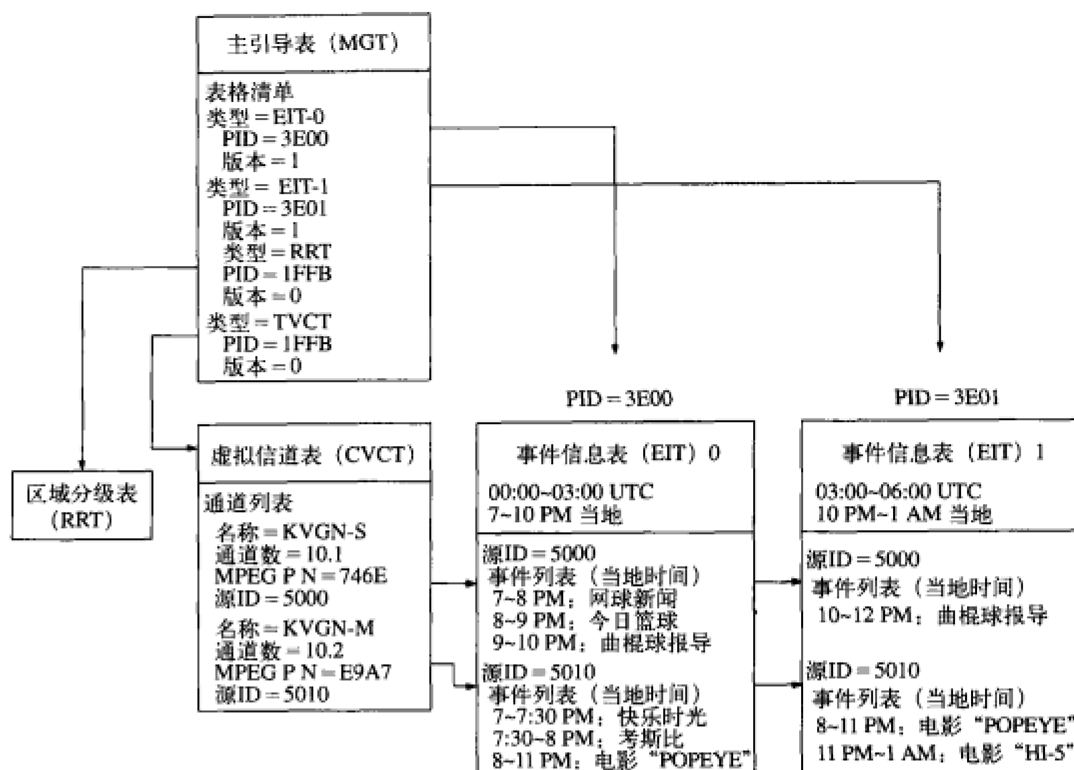


图15-1 ATSC PSIP表的关系

2. 主引导表（MGT）

该表包含了其他表的一般信息。它定义了表的大小、版本号以及包标识符（PID）。

3. 分级区域表（RRT）

这一表传输的是通常称为V-chip的分级系统。

4. 系统时间表（STT）

系统时间表是日期、时间的参考。接收端用它来维持正确的当地时间。

5. 地面虚拟信道表（TVCT）

除了线缆VCT（CVCT）和卫星VCT（SVCT）之外，还有地面虚拟信道表，又称为VCT。它包含一个列表，该表中列出了传输流中所有可用的或将会可用的信道以及它们的特征。VCT还可以包含其他传输流中广播的模拟信道和数字信道。

每个信道的特征包括最大/最小信道数、缩写、唯一标识每个基站的传输/发射系统ID（TSID）等。服务位置描述符列出了视频、音频、数据和其他相关基本流的PID。

15.3.2 可选表

1. 扩展文本表（ETT）

对于文本信息，可以有多个ETT，每个ETT的PID由MGT定义。表中的信息用来描述通道信息、即将播放的视频（coming attractions）、电影描述等。

768

2. 定向频道转换表（DCCT）

DCCT中包含在广播商设定参数时完成频道转换所需要的信息。频道转换可能是无条件的，

也可能是依据观众所设定的准则进行的。

3. 定向频道转换选择编码表（DCCSCT）

通过一些定向频道转换请求，DCCSCT允许下载广播节目分类表。

15.3.3 描述符

和MPEG-2非常相似，ATSC采用描述符来添加新功能。除了各种MPEG-2描述符之外，在PMT、一个或多个PSIP表内可以包含一个或多个ATSC专有描述符，以扩展表内数据。解码器不能识别的描述符必须丢弃。这使得添加新功能不影响不能识别和处理描述符的接收端。

769 ~ 770

1. AC-3音频流描述符

该ATSC描述符表示存在Dolby® Digital音频或Dolby® Digital Plus音频，第13章对此进行了详细讨论。

2. ATSC CA描述符

该ATSC描述符的语法和MPEG-2 CA描述符的语法几乎完全一样。

3. ATSC私有信息描述符

该ATSC描述符提供了一种携带私有信息的方式，第13章对此进行了详细讨论。同时，在一个描述符中又可以包含一个或多个描述符。

4. 组件名称描述符

该ATSC描述符为服务的每个组件定义了一个可变长的基于内容的名称。第13章对此进行了详细讨论。

5. 内容分级描述符

该ATSC描述符定义了每个特定节目的级别，第13章对此进行了详细讨论。

6. 内容标识描述符

该ATSC描述符用于唯一标识ATSC传输中的内容。

7. DCC到达请求描述符

一旦转换到一个新频道，该ATSC描述符就为接收端执行相应的动作提供指令：显示至少10秒的内容，或者，如果观众发出了“继续”、“确定”或类似的命令，显示的时间更短；显示内容的时间长度不确定，或者，直到观众发出了“继续”、“确定”或类似的命令。

8. DCC离开请求描述符

当离开某个频道之前，该ATSC描述符就为接收端执行相应的动作提供指令：

取消所有未完成的操作，直接切换频道；

显示至少10秒的内容，或者，如果观众发出了“继续”、“确定”或类似的命令，显示的时间更短；

显示内容的时间长度不确定，直到观众发出了“继续”、“确定”或类似的命令。

9. 增强信令描述符

该ATSC描述符标识了节目单元的地面广播传输方式，第13章对此进行了详细讨论。

10. 扩展通道名描述符

该ATSC描述符为虚拟通道提供了一个可变长的频道名称。

11. 流派描述符

该ATSC描述符提供了流派、节目类型或事件的分类信息，并且对于给定EIT事件，它可能会出现在descriptor()循环中。它引用分类流派编码赋值表（Categorical Genre Code Assignments Table）中的项，还要引用DCC选择码提供的表格扩展。

12. 再发行控制描述符

该ATSC描述符给出了节目内容版权所有者所持有的所有再发行控制信息。第13章对此进行了详细讨论。

771

13. 服务位置描述符

该ATSC描述符定义了流类型、PID和每个基本流的语言码。在每个有效通道的VCT中都有该描述符。

14. SRM参考描述符

该ATSC描述符是第13章中讨论的MPEG-2 CA描述符的一个特定实现。用它来指明是否为系统更新能力信息表提供了系统更新能力信息。

15. 时间平移服务描述符

该ATSC描述符将一个虚拟通道链接到最多20个传播同样节目的虚拟通道，但是不同通道之间节目的时间发生了平移。典型的应用是准视频点播（NVOD）服务。

15.4 E-VSB

E-VSB又称为"增强8-VSB"，它使得ATSC广播可以包含一个更低比特率的二级节目，该节目比弱信号条件下的典型HDTV节目具有更好的健壮性。如果干扰降低了原有HDTV的信号强度，这时接收端可以切换到复用于同一传输流中的、更健壮的同一节目流的SDTV信号。例如，在19Mbit/s信道内，14Mbit/s可以用于HD节目，4Mbit/s用于健壮的SD节目，而1Mbit/s用于控制信息。

当E-VSB不用于后向音频特征时，它可以用来提供附加服务，例如增强音频、附加只包含音频的服务或者元数据，以控制两个音频流的混合。

15.4.1 音频能力

对于增强服务，音频压缩算法采用的是Dolby® Digital音频或Dolby® Digital Plus音频，并且支持1~5.1声道，采样率为48kHz，最大比特率为448kbit/s。允许采用杜比数字Dolby® Digital音频，但是采用Dolby® Digital Plus音频可以为主节目维持尽可能高的比特率。

音频服务类型限制为CM、VI、HI和C。每种音频服务必须包含一个完整的音频节目（对话、音乐和音效）。

当E-VSB不用于后向音频特征时，可以采用5.1以上的声道或高于48kHz的采样率。

15.4.2 视频能力

增强服务中的视频压缩算法倾向于采用MPEG-4.10（H.264）甚至HP@L4.0。允许采用MPEG-2，但是采用MPEG-4.10（H.264）可以为主节目维持尽可能高的比特率。

15.4.3 节目和系统信息协议

节目和系统信息协议（PSIP-E）是采用E-VSB模式传输的节目和系统信息。

PAT-E：节目关联表。语法和PAT相同。

772

PMT-E：节目映射表。语法和PMT相同。

15.5 数据广播

ATSC数据广播标准描述了在MPEG-2传输流中传输数据的各种方法。它可以应用于很多领域，如：传递声明数据（HTML码）；传递程序数据（Java码）；传递软件和图像，传递MPEG-4.2、

MPEG-4.10 (H.264)或SMPTE 421M (VC-1)视频流，混合MPEG-2视频或MP3音频文件。

标准所定义的关键内容有：数据服务声明，数据传递模式，如数据管道、数据流、可寻址片段和数据下载，应用标识，MPEG-2系统工具，协议。

数据服务必须包含在虚拟通道中，并且每个虚拟通道最多只能有一个数据服务。一个数据服务可以包含多个应用，每个应用可以包含多个数据单元。

第13章同样对数据广播进行了讨论。

15.5.1 数据服务声明

数据广播采用并扩展了PSIP，目的是在广播流中声明并找到数据服务。数据服务通过在PSIP EIT或数据事件表（DET）中的事件进行声明。

数据服务声明的附加表包括以下几项（在此不讨论数据服务的附加描述符）。

1. 数据事件表（DET）

数据事件表最多有128个，DET-0~DET-127。每个数据事件表描述了虚拟通道中的数据服务的信息（标题、开始时间等）。每个DET的有效时间是3小时。所需要DET表的最少个数是4 (DET-0~DET-3)。DET中发生的任何变化都将导致MGT也发生变化。

2. 扩展文本表（ETT）

ETT用于描述数据事件的详细信息。其语法和用于AV服务的ETT基本类似。

3. 长期服务表（LTST）

LTST用于预声明在EIT/DET所能支持的时间范围之外发生的数据事件。

15.5.2 服务描述框架

由于协议封装的形式可能多种多样，因此，在每个数据流中都需要表明采用了哪种封装形式。虽然可以采用PMT来实现这一功能，但这种方法的可扩展性差，只用于简单情况。*数据描述框架*可以提供一种可扩展的实现框架。

服务描述框架（SDF）所需要的附加表包括以下几项。 773

1. 数据服务表（DST）

DST描述由一个或多个接收端应用构成的数据服务，它同时提供相关信息以允许数据接收端将应用与其所消耗的数据相关联。

2. 网络资源表（NRT）

NRT提供了当前MPEG-2程序或传输流之外的所有网络资源的列表。

数据服务可以用NRT来获得除了在VCT的*服务位置描述符*内发布的数据包之外的数据包或数据报。这包括同一传输流中所包含的其他节目的数据基本流，其他传输流中的数据基本流，采用诸如IP的其他协议的双向通信信道。

15.5.3 触发器（同步和异步）

有了对同步和异步触发器的支持，通过从数据元素的传输中解耦出时序信号就可以实现数据模块的同步传输。它同时允许向接收端传递事件，包括应用定义的事件。在下载协议中触发器作为一个数据模块传递。

触发器携带指向将要激发的目标的指针、指向需要处理用于显示的预加载数据的应用的指针或额外的自包含用户数据。

当接收端的STC的90kHz部分和触发器定义的PTS值匹配时，将激活同步触发器。一旦触发器

解码完成就会激活异步触发器。

15.5.4 软件下载数据服务

软件下载数据服务（SDDS）定义了软件的传输。它是建立在数据广播标准中定义的数据服务传递机制的基础上的。SDDS支持软件的传输，包括更新固件、中间件、应用程序、操作系统等。

15.5.5 传输流文件系统

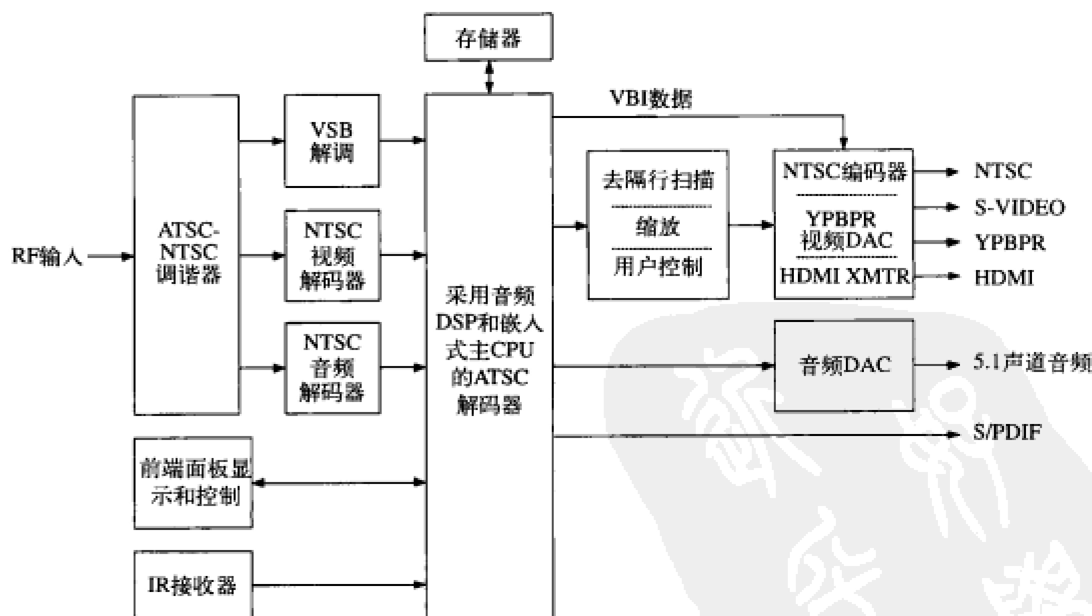
传输流文件系统（TSFS）定义了层次命名空间、目录和文件的传输。它是建立在数据广播标准中定义的数据服务传递机制基础上的。TSFS支持采用MPEG-2.6 DSM-CC数据混合协议传输目录、文件和服务网关目标。

15.6 应用框图

图15-2示出了典型的ATSC接收器机顶盒框图。机顶盒通常要求能够同时输出高清和标清版本的节目。

图15-3示出了典型的ATSC数字电视的框图。为了支持PIP（Picture-In-Picture），通常要求数字电视能够同时解码两个节目。

774



775

图15-2 典型的ATSC接收端机顶盒框图

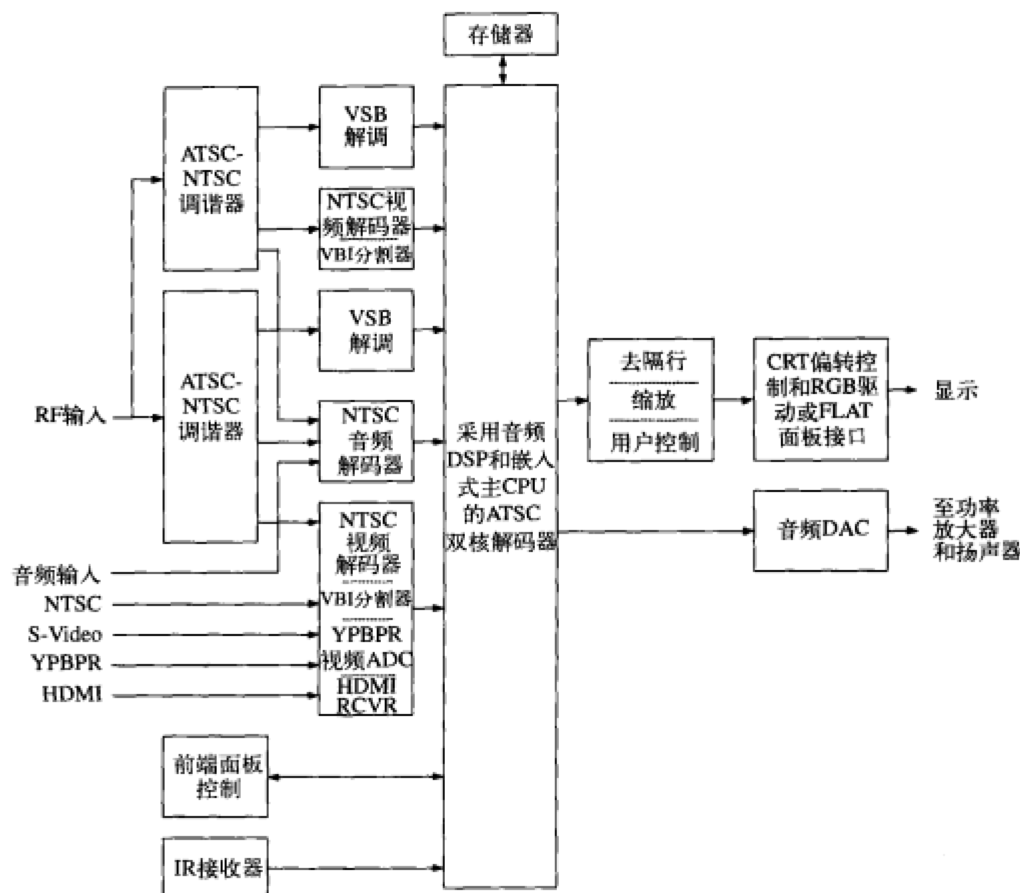


图15-3 典型的ATSC数字电视框图

776

参考文献

1. ATSC A/52B, *Digital Audio Compression Standard (AC-3, E-AC-3)*, June 2005.
2. ATSC A/53E, *ATSC Digital Television Standard*, September 2006.
3. ATSC A/54A, *Recommended Practice: Guide to the Use of the ATSC Digital Television Standard*, December 2003.
4. ATSC A/57A, *ATSC Standard: Content Identification and Labeling for ATSC Transport*, July 2003.
5. ATSC A/58, *Harmonization With DVB SI in the Use of the ATSC Digital Television Standard*, September 1996.
6. ATSC A/64A, *Transmission Measurement and Compliance for Digital Television*, May 2000.
7. ATSC A/65C, *ATSC Standard: Program and System Information Protocol for Terrestrial Broadcast and Cable*, January 2006.
8. ATSC A/69, *ATSC Recommended Practice: Program and System Information Protocol Implementation Guidelines for Broadcasters*, June 2002.
9. ATSC A/70A, *ATSC Standard: Conditional Access System for Terrestrial Broadcast*, July 2004.
10. ATSC A/76A, *ATSC Standard: Programming Metadata Communication Protocol Standard*, September 2006.
11. ATSC A/80, *Modulation and Coding Requirements for Digital TV (DTV) Applications Over Satellite*, July 1999.
12. ATSC A/81, *ATSC Direct-to-Home Satellite Broadcast Standard*, July 2003.
13. ATSC A/90, *ATSC Data Broadcast Standard*, July 2000.
14. ATSC A/91, *Recommended Practice: Implementation Guidelines for the ATSC Data Broadcast Standard*, June 2001.
15. ATSC A/92, *ATSC Standard: Delivery of IP Multicast Sessions over ATSC Data Broadcast*, January 2002.
16. ATSC A/93, *ATSC Standard: Synchronized/Asynchronous Trigger*, April 2002.
17. ATSC A/94, *ATSC Standard: Data Application Reference Model*, August 2002.
18. ATSC A/95, *ATSC Standard: Transport Stream File System Standard*, February 2003.
19. ATSC A/96, *ATSC Standard: Interaction Channel Protocols*, February 2004.
20. ATSC A/97, *ATSC Standard: Software Download Data Service*, November 2004.

777

第16章　OpenCable™数字电视

OpenCable™是美国的一个数字有线标准，其目的是为不同的硬件和软件开发商提供互操作性。该标准的一个子集逐渐被数字电视采纳。

其他三个主要的数字电视标准是ATSC（Advanced Television Systems Committee）、DVB（Digital Video Broadcast）和ISDB（Integrated Services Digital Broadcasting）。这些标准的基本音频和视频能力非常相似，主要区别在于RF调制机制和非音/视频服务的定义。表16-1对OpenCable™标准进行了总结。

表16-1　OpenCable™标准总结

参　数	OpenCable™
视频压缩	MPEG-2
音频压缩	Dolby® Digital音频
复用	MPEG-2传输流
调制	QAM
信道带宽	6MHz

OpenCable™基于下列及其他ATSC和SCTE标准：

A/52——数字音频压缩（AC-3和EAC-3）标准
A/53——ATSC数字电视标准
A/65——地面广播和有线的节目与系统信息协议
A/90——ATSC数据广播标准
SCTE 07——有线电视数字传输标准
SCTE 18——有线紧急报警信息
SCTE 20——传递剧终字幕和非实时采样视频的方法
SCTE 26——具有版权保护的家庭数字网络接口定义
SCTE 40——数字有线网络接口标准
SCTE 43——有线电视数字视频系统特征标准
SCTE 54——有线电视数字视频服务多路复用和传输系统
SCTE 55——数字广播传输系统，带外传输
SCTE 65——用于数字有线电视的带外传输服务信息
SCTE 80——包含带外声明的带内数据广播标准

OpenCable™接收器在数字有线网络上使用下列4种通信信道。

- 6MHz NTSC模拟信道。它们通常位于54~450MHz范围内。每个信道携带一个节目。
- 6MHz前端应用传输（FAT）信道。它们是通过MPEG-2传输流传输内容的。它们采用QAM编码算法并且通常位于450~864MHz范围内。每个通道可以携带多个节目。
- 带外（OOB）前端数据信道（FDC）。它们采用QPSK调制，并且通常位于70~130MHz范围内，位于6MHz NTSC模拟和FAT信道之间。SCTE 55-1和SCTE 55-2是两种可选实现方案。
- 带外（OOB）反向数据信道（RDC）。它们采用QPSK调制，并且通常位于5~42MHz范围内。SCTE 55-1、SCTE 55-2和DOCSIS提供了3种可选的实现方案。

OpenCable™接收器通过调谐到有线电视连接上可用的多个6MHz的信道之一来获得数据内容。当选择的信道是模拟信道时，信号采用NTSC音/视频/VBI解码器进行处理。

当所选择的信道是数字信道时，信号采用QAM解调器进行处理，然后采用CableCARD™对

内容进行解扰（条件接收解扰）。条件接收解扰针对特定的有线系统，并且通常是专有的。之后，CableCARD™采用一个常规算法将内容重新加扰，并将其传递到MPEG-2解码器。多数据流CableCARD™可以同时传送多达6个不同的信道，从而可以进行画中画和DVR（数字视频录制）服务。

当没有插入CableCARD™时，数字调谐器的QAM解调器的输出直接导入MPEG-2解码器。然而，加密的部分是不可见的。

CableCARD™接收器通过调谐到OOB FDC信道获得控制信息和其他数据。采用专用调谐器，接收器可以连续从OOB FDC信道接收信息。接收到的信息同时被传递到CableCARD™和MPEG-2解码器进行处理。

779

双向OpenCable™接收器同样可以采用OOB FDC传输数据。

OpenCable™标准采用MPEG-2传输流在一个单6MHz FAT通道上传输压缩的视频信息、压缩音频信息和附加数据。多视频流、多音频流和数据可以存在于MPEG-2传输流中。

MPEG-2传输流的比特率固定为约27Mbit/s（64-QAM调制）、33.8Mbit/s（256-QAM）或44.3Mbit/s（1024-QAM）。

可用比特率可以有许多种使用方式，具体采用何种方式需要在所提供的节目数和视频质量及分辨力之间进行权衡。例如，如果采用MPEG-2视频、统计复用和256-QAM，那么：

（4）HDTV节目

（2）HDTV节目＋（6）SDTV节目＋数据

（18）SDTV节目

16.1 视频能力

数字视频压缩采用MPEG-2实现，并且它要求和ATSC一样。但是对于一些MPEG-2参数，有一些附加的要求，这一点如同在MPEG-2一章所讨论的一样。标准的定义中同时增加了MPEG-4.10（H.264）甚至HP@L4.0。

虽然只要不超过最大比特率，就可以采用任意分辨率，但通常只采用几个标准化的分辨率，如表16-2所列。对于大多数分辨率允许采用隔行或逐行两种方式。

与标准兼容的接收器必须能够调谐到模拟NTSC信号并对其解码，这一点在第8章中进行了讨论。

16.2 音频能力

数字音频压缩采用Dolby® Digital音频实现，并且它的要求和ATSC的一样。

与标准兼容的接收器必须能够解码模拟NTSC信号中的音频数据。NTSC音频标准在第8章中进行了讨论。

16.3 带内系统信息

在MPEG-2传输流中有足够的带宽以支持多个低带宽、非电视服务，如节目导航、隐藏字幕、天气预报、股市指数、热点新闻、软件下载、按次付费信息等。附加非电视服务的数量（虚拟通道）很容易达到十个或者更多。此外，服务的种类也会不断变化。

780

表16-2　OpenCable™视频常见有效分辨率

有效分辨率（Y）	SDTV或HDTV	帧率（0＝逐行，1＝隔行）			
		23.976p 24p	29.97i 30i	29.97p 30p	59.94p 60p
480×480	SDTV	×	×	×	×
528×480		×	×	×	×
544×480		×	×	×	×
640×480		×	×	×	×
704×480		×	×	×	×
1280×720	HDTV	×		×	×
960×1080		×	×	×	
1280×1080		×	×	×	
1440×1080		×	×	×	
1920×1080		×	×	×	

定义系统信息（SI）是为了以一种灵活、一致的方式支持这些非电视服务。SI是一个层次式关联表集合（见图16-1和表16-3），这些表用于扩展MPEG-2 PSI表。SI描述了在一个特定的MPEG-2传输流中携带的所有虚拟通道的信息。此外，针对模拟广播信道的信息也可以包涵在其中。

对于带内SI，OpenCable™绝大部分沿用了ATSC PSIP标准，只是进行了一些扩展。

16.3.1　所需表

1. 有线虚拟通道表（CVCT）

这一表格列出了传输流中所有可用或即将可用的信道及其特性。同时，它也可能包括其他传输流中的模拟信道和数字信道。

每个信道的特性包括最大/最小信道数、简写、唯一标识每个基站的传送/传输系统ID（TSID）、载波频率、调制模式等。服务位置描述符用来列出视频、音频、数据和其他相关基本流的PID。

781

ATSC也采用这种表格，在ATSC中这种表格称为地面虚拟通道表（TVCT）。

2. 紧急报警（EA）表

该表提供了一种标识方式，使得有线电视操作员可以向数字机顶盒、数字电视接收器、数字VCR等设备发送紧急信号。这些设备必须能够存储任何EA事件以备后用。开始时间和持续时间信息用来删除过期的事件。

典型情况下，来源于与有线集线器相同地理区域的地面数字广播源的传输流同样在其广播中提供紧急事件信息。

3. 事件信息表（EIT）

事件信息表总共有128个，包括EIT0~EIT127，每个EIT描述在CVCT中列出的每个虚拟通道的事件或电视节目。每个EIT的有效时间是3小时。由于总共有128个EIT，未来16天内的节目都可以提前预告。前4个EIT是必需的（建议的使用前24个）。

EIT提供的信息包括开始时间、持续时间、标题、指向用于描述事件的可选文本的指针、分级数据（advisory data）、字幕服务数据、音频服务描述符等。ATSC也采用这种表。

4. 主引导表（MGT）

该表包含了其他表的一般信息。它定义了表的大小、版本数以及包标识符（PID）。ATSC也采用这种表。

5. 分级区域表（RRT）

这一表传输的是通常称为V-chip的分级系统。ATSC也采用这种表。

6. 系统时间表（STT）

系统时间表是日期、时间的参考。接收端用它来维持正确的当地时间。ATSC也采用这种表。

16.3.2 可选表

1. 定向信道变化表（DCCT）

DCCT包含在设定的广播时间内实现信道变化所需要的信息。信道变化可能是无条件的，也可能是基于观众设定的要求。ATSC也采用这种表。

2. 定向信道变化选择码表（DCCSCT）

DCCSCT允许通过一些定向信道变化请求下载广播节目分类表。ATSC也采用这种表。

3. 扩展文本表（ETT）

对于文本信息，可以有多个ETT，每个ETT的PID由MGT定义。表中的信息用来描述通道信息、后续的有趣事物、电影描述等。ATSC也采用这种表。

782

表16-3 OpenCable™的带内SI表、描述符和描述符位置

描述符	描述符标识	地面广播表									
		PMT	MGT	CVCT	RRT	EIT	ETT	STT	DCCT	DCCSCT	CAT
PID		每PAT	0x1FFB	0x1FFB	0x1FFB	每MGT	每MGT	0x1FFB	0x1FFB	0x1FFB	0x0001
Table_ID		0x02	0xC7	0xC9	0xCA	0xCB	0xCC	0xCD	0xD3	0xD4	0x80, 0x81 (ECM) 0x82~0x8F (EMM)
重复率		400ms	150ms	400ms	1min	0.5s	1min	10s	400ms	1h	
AC-3音频流	1000 0001	M									
ATSC CA	1000 1000			O		O					
ATSC私有信息*	1010 1101										
CA	0000 1001	M									M
字幕服务	1000 0110	M				M					
组件名称	1010 0011	M									
内容分级	1000 0111	M				M					
DCC到达请求	1010 1001								M		
DCC离开请求	1010 1000								M		
扩展通道名称	1010 0000			M							
扩展视频	1000 0011	M									
帧率	1000 0010	M									
MAC地址列表	1010 1100	M									
重分配控制	1010 1010	M				M					
服务位置	1010 0001			M							
填充信息*	1000 0000										
时移服务	1010 0010			M							

注：1. PMT：MPEG-2节目映射表。CAT：MPEG-2条件接收表。

2. M＝当存在时，该表中需要；O＝在该表中也可能存在，*＝没有要求。

783

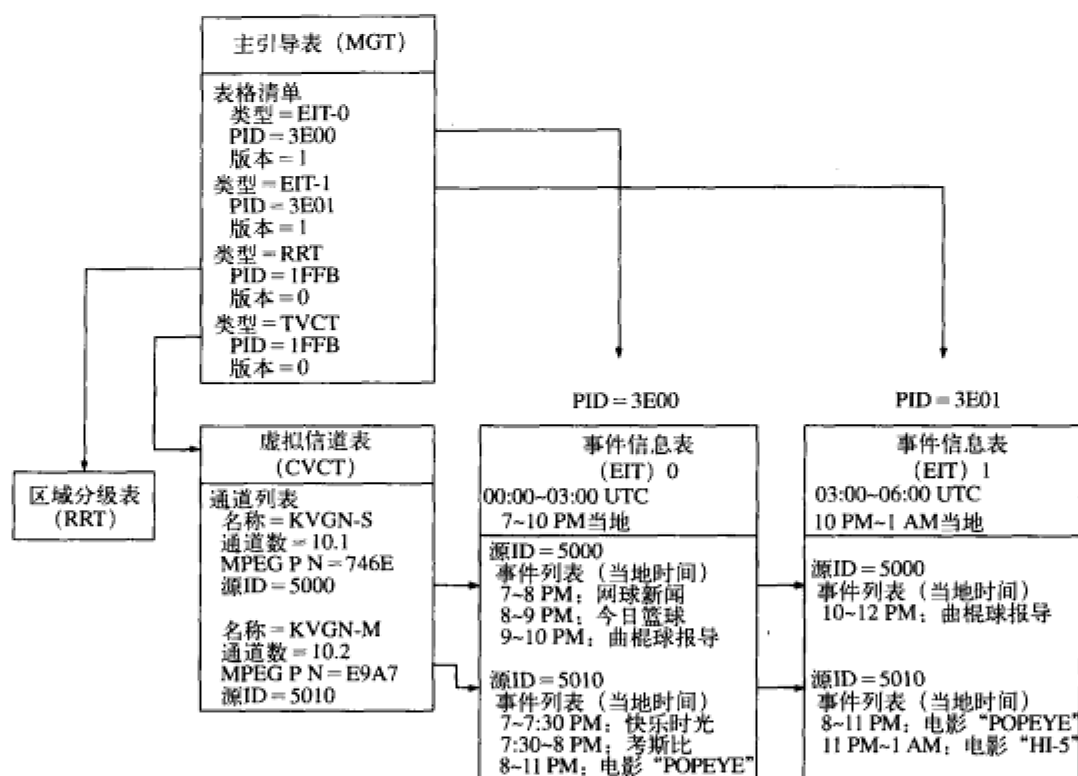


图16-1　OpenCable™的带内SI表的关系

16.3.3　描述符

和MPEG-2非常相似，OpenCable™采用描述符来添加新功能。除了各种MPEG-2描述符之外，在PMT、一个或多个SI表（见表16-3）内可以包含一个或多个OpenCable™专有描述符，以扩展表内数据。解码器不能识别的描述符必须丢弃。这使得添加新功能不影响不能识别和处理描述符的接收端。

1. AC-3音频流描述符

这一OpenCable™描述符表示存在Dolby® Digital音频或Dolby® Digital Plus音频，第13章对此进行了详细讨论。ATSC也采用了这种描述符。

2. OpenCable™私有信息描述符

这一OpenCable™描述符提供了一种携带私有信息的方式，第13章对此进行了详细讨论。同时，在一个描述符中又可以包含一个或多个描述符。ATSC也采用这种描述符。

784

3. 组件名称描述符

这两个OpenCable™描述符为服务的每个组件定义一个可变长的基于内容的名称。第13章对此进行了详细讨论。ATSC也采用了其中一个描述符。

4. 内容分级描述符

这一OpenCable™描述符定义了每个特定节目的级别，第13章对此进行了详细讨论。ATSC也采用了这种描述符。

5. DCC到达请求描述符

这一OpenCable™描述符定义了到达新的变化后的通道后接收端所进行的操作：显示至少10秒的内容，或者，如果观众发出了“继续”、“确定”或类似的命令，显示的时间更短；显示内容的时间长度不确定，直到观众发出了“继续”、“确定”或类似的命令。ATSC也使用这种描述符。

6. DCC离开请求描述符

这一OpenCable™描述符定义了接收端离开某一通道之前所要进行的操作：取消所有未完成的操作，直接切换通道；显示至少10秒的内容，或者，如果观众发出了“继续”、“确定”或类似的命令，显示的时间更短；显示内容的时间长度不确定，直到观众发出了“继续”、“确定”或类似的命令。ATSC也采用了这种描述符。

7. 扩展通道名称描述符

这一OpenCable™描述符为虚拟通道提供了一个可变长的通道名称。ATSC也采用了这种描述符。

8. 扩展视频描述符

这一OpenCable™描述符标识了在处理过程中可能需要的信息，第13章对此进行了详细讨论。

9. 帧率描述符

这一OpenCable™描述符标识视频图像的帧率，第13章对此进行了详细讨论。

10. MAC地址列表描述符

在MPEG-2传输流中实现IP多点传送时，使用这一OpenCable™描述符。第13章对此进行了详细讨论。

11. 再发行控制描述符

这一OpenCable™描述符传递节目内容的版权持有者所持有的任何再发行控制信息。第13章对此进行了详细讨论。ATSC也采用了这种描述符。

12. 服务位置描述符

这一OpenCable™描述符定义了流类型、PID和每个基本流的语言码。它包含在每个活跃通道的CVCT中。ATSC也采用了这种描述符。

13. 时移服务描述符

这一OpenCable™描述符将一个虚拟通道链接到最多20个传播同样节目的虚拟通道，但是节目的时间移动了。其典型的应用是准视频点播（NVOD）服务。ATSC也采用了这种描述符。

785

16.4 带外系统信息

带外系统信息（SI）数据可以在带外传输。CableCARD™将一个兼容MPEG-2传输流中可能存在也可能不存在的OOB SI数据转换为兼容的表区（table section），每个表区赋予一个适当的PID值。

根据不同的网络配置，OOB消息可以在OOB FDC、OOB RDC通道或DOCSIS通道上实现。CableCARD™通知接收器采用哪个系统。

16.4.1 表

总共定义了6个档次（见表16-4）来表示所需和可选的表。与这些表保持一致是兼容性的要求之一。

1. 汇总事件信息表（AEIT）

该表包含了支持EPG的事件名称和调度信息。为了减少用于SI数据的PID值的总体数量，

AEIT格式允许不同时间周期调用的表区与常用PID值相关联。

2. 汇总扩展文本表（AETT）

该表包含了可以用于传递对事件的详细描述的扩展文本信息。ETM是一个多串数据结构，因此可以用几种不同的语言传递对事件的描述。

3. 紧急报警（EA）表

该表提供了一种信令发送方式，使得有线电视操作员可以向数字机顶盒、数字电视接收器、数字VCR等设备发送紧急信号。这些设备必须能够存储任何EA事件以备后用。开始时间和持续时间信息用来删除过期的事件。

4. 长型虚拟信道表（L-VCT）

该表是采用MPEG-2私有部分传输的CVCT。

5. 主引导表（MGT）

该表包含了其他表的一般信息。它定义表的大小、版本号以及包标识符（PID）。ATSC也使用这种表。

6. 网络信息表（NIT）

该表将一组传输流集中在一起，为接收器提供了调谐信息。

7. 网络文本表（NTT）

该表包含系统级的多语言文本串。

8. 分级区域表（RRT）

这一表传输的是通常称为V-chip的分级系统。ATSC也采用这种表。

9. 短型虚拟信道表（S-VCT）

786 该表包含部分虚拟信道映射（VCM）、详细通道映射（DCM）和反通道映射（ICM）信息。

表16-4 不同的档次中所采用的表

表	表ID	档次1 基本	档次2 修改检测	档次3 来源咨询	档次4 标准EPG数据	档次5 组合	档次6 仅SI
NIT	0xC2						
载波定义子表		M	M	M	M	M	
调制模式子表		M	M	M	M	M	
NTT							
源名称子表	0xC3	O	O	O	M	M	
短型VCT	0xC4						
虚拟信道映射		M	M	M	M	M	
详细信道映射		M	M	M	M	M	
反信道映射		O	O	O	O	O	
STT	0xC5	M	M	M	M	M	M
MGT	0xC7			M	M	M	M
RRT	0xCA			M	M	M	M
L-VCT	0xC9					M	M
AEIT	0xD6				M	M	M
AETT	0xD7				O	O	O

787

注：M＝当存在时，该表中需要，O＝在该表中也可能存在。

10. 系统时间表（STT）

系统时间表是日期、时间的参考。接收端用它来维持正确的当地时间。ATSC也采用这种表。

16.4.2 描述符

表16-5和表16-6分别描述了不同的档次和表中所采用的描述符。

表16-5 不同的档次中所采用的描述符

描述符	标识	档次1 基线	档次2 修改检测	档次3 来源咨询	档次4 标准EPG数据	档次5 组合	档次6 仅SI
AC-3音频	0x81				O	O	O
字幕服务	0x86				O	O	O
信道特性	0x95				O	O	
组件名称	0xA3				O	O	O
内容分级	0x87			M	M	M	M
夏令时	0x96			O	M	M	M
扩展信道名	0xA0					O	O
修改检测	0x93		M	M	M	M	
时移服务	0xA2					O	O
两部分信道编号	0x94				O	O	

注：M＝当存在时，该表中需要；O＝在该表中也可能存在。

表16-6 不同的表中所采用的描述符

描述符	标识	表								
		PMT	NIT	NTT	S-VCT	STT	MGT	L-VCT	RRT	AEIT
PID		每 PAT	0x1 FFC	0x1 FFC	0x1 FFC	0x1 FFC	0x1 FFC	0x1 FFC	0x1 FFC	每 MGT
Table_ID		0x02	0xC2	0xC3	0xC4	0xC5	0xC7	0xC9	0xCA	0xD6
AC-3音频	0x81	×								×
字幕服务	0x86	×								×
信道特性	0x95				×					
组件名称	0xA3	×								
内容分级	0x87	×								×
夏令时	0x96					×				
扩展信道名称	0xA0							×		
修改检测	0x93		×	×	×					
时移服务	0xA2							×		
两部分信道编号	0x94				×					

1. AC-3音频流描述符

这一描述符表示存在Dolby® Digital音频，第13章对此进行了详细讨论。ATSC也采用了这种描述符。

2. 信道特性描述符

该描述符使得接收器能够识别不同的信道特性。否则，接收器必须调谐到信道并自动提取信道特征。

3. 组件名称描述符

这两个OpenCable™描述符为服务的每个组件定义了一个可变长的基于内容的名称。第13章对此进行了详细讨论。ATSC也采用了其中一个描述符。

4. 内容分级描述符

这一描述符定义了每个特定节目的级别，第13章对此进行了详细讨论。ATSC也采用了这种

788 描述符。

5. 夏令时描述符

该描述符表示当前状态下是否检测夏令时，以及夏令时发生在什么时间、什么日期。

接收器不能认为没有该描述符就意味着夏令时当前没有生效。

6. 扩展通道名称描述符

这一描述符为虚拟通道提供了一个可变长的通道名称。ATSC也采用了这种描述符。

7. 修改检测描述符

这一描述符表示描述符所在的表区（table section）中是否存在新的信息。为了降低处理开销，该描述符应该是列表的第一个描述符。

8. 时移服务描述符

这一描述符将一个虚拟通道链接到最多20个传播同样节目的虚拟通道，但是节目的时间移动了。其典型的应用是准视频点播（NVOD）服务。ATSC也采用了这种描述符。

9. 两部分信道编号描述符

789 该描述符可以用于将一个两部分（two-part）信道编号（即10-2）与任何虚拟信道关联起来。

16.5　带内数据广播

OpenCable™带内数据广播标准描述了在MPEG-2传输流中传输数据的各种方法。它可以应用于很多应用领域，如：

传递声明数据（HTML码）

传递程序数据（Java码）

传递软件和图像

传递MPEG-4.2、MPEG-4.10（H.264）或SMPTE 421M（VC-1）视频流

混合MPEG-2视频或MP3音频文件

标准所定义的关键内容有：

数据服务声明

数据传递模式，如数据管道、数据流、可寻址片段和数据下载

应用标识

MPEG-2系统工具

协议

数据服务必须包含在虚拟通道中，并且每个虚拟通道最多只能有一个数据服务。一个数据服务可以包含多个应用，每个应用可以包含多个数据单元。

对于带内数据广播，OpenCable™基本沿用了ATSC数据广播标准。主要区别在于OpenCable™增加了带外声明。

第13章同样对数据广播进行了讨论。

16.5.1　数据服务声明

数据广播采用并扩展了SI，以便在广播流中声明并找到数据服务。数据服务通过EIT或数据事件表（DET）中的事件进行声明。

数据服务声明的附加表包括以下几项（在此不讨论数据服务的附加描述符）。

1. 数据事件表（DET）

数据事件表最多有128个，从DET-0到DET-127，每个数据事件表描述了虚拟通道中数据服务的信息（标题、开始时间等）。每个DET的有效时间是3小时。所需要的最少的DET表的个数是4

个（DET-0~DET-3）。DET中发生的任何变化将使MGT也发生变化。ATSC也支持这种表。

2. 扩展文本表（ETT）

ETT用于提供数据事件的详细描述。其语法和用于AV服务的ETT基本类似。ATSC也支持这种表。

3. 长期服务表（LTST）

LTST用于预声明在EIT/DET所能支持的时间范围之外发生的数据事件。ATSC也支持这种表。 790

16.5.2 服务描述框架

由于可能的协议封装的形式多种多样，因此，在每个数据流中需要表明采用了哪种封装形式。虽然可以采用PMT来实现这一功能，但这种方法的可扩展性差，只用于简单情况。*数据描述框架*可以提供一种可扩展的实现框架。

服务描述框架（SDF）所需要的附加表包括以下几项（这里不讨论针对数据服务的附加描述符）。

1. 汇总数据事件表（ADET）

ADET最多有128个（ADET-0~ADET-127）。它们包含了实现电子节目导航（EPG）所需要的事件名称和调度信息。采用ADET的目的有以下两个。

- 在不包含任何音频－虚拟事件的虚拟通道中声明一个数据服务。
- 在虚拟通道中，允许独立声明音频/视频/数据服务或音频/数据服务的数据服务部分。

由于CableCARD™只能支持小数目的并行数据流（每个关联到一个PID值），传输格式允许ADET表区共享一个通用PID值。每个ADET的有效时间是3小时。所需的ADET最小数目为4（ADET0~ADET3）。ADET的任何变化都会使得MGT发生变化。

ADET分区表可以通过带内（MPEG-2传输流）或带外的方式传输到接收器。

2. 数据服务表（DST）

DST描述由一个或多个接收端应用构成的数据服务，它同时提供相关信息以允许数据接收端将应用与其所消耗的数据相关联。ATSC也支持这种表。

3. 网络资源表（NRT）

NRT提供了当前MPEG-2程序或传输流之外的所有网络资源的列表。

数据服务可以用NRT来获得除了在CVCT的服务位置描述符内发布的数据包之外的数据包或数据报。这包括同一传输流中包含的另一个节目的数据基本流，其他传输流中的数据基本流，采用诸如IP的其他协议的双向通信信道。ATSC也支持这种表。

16.6 条件接收

OpenCable™中采用的条件接收策略与多密技术和第17章中讨论的DVB通用接口类似。在OpenCable™中，CA模块称为CableCARD™，并且它也基于EIA679NRSS-B接口（PCMCIA或PC卡形式的接口）。针对DVB方案的主要的两点扩充包括同时支持多达6个数据流的能力（要求具有多数据流CableCARD™）和位于CableCARD™输出与MPEG-2解码输入之间的DFAST加密接口。

为了简化器件，可以采用可下载的条件接收系统（DCAS），从而可以避免使用CableCARD™。 791

16.7 相关技术

除了OpenCable™之外，CableLabs®已经开发出了并且还在继续开发各种与有线相关的标准。

❑ DOCSIS®

DOCSIS®（基于有线服务接口规范的数据传输）定义了有线调制解调器的接口要求。

❑ PacketCable™

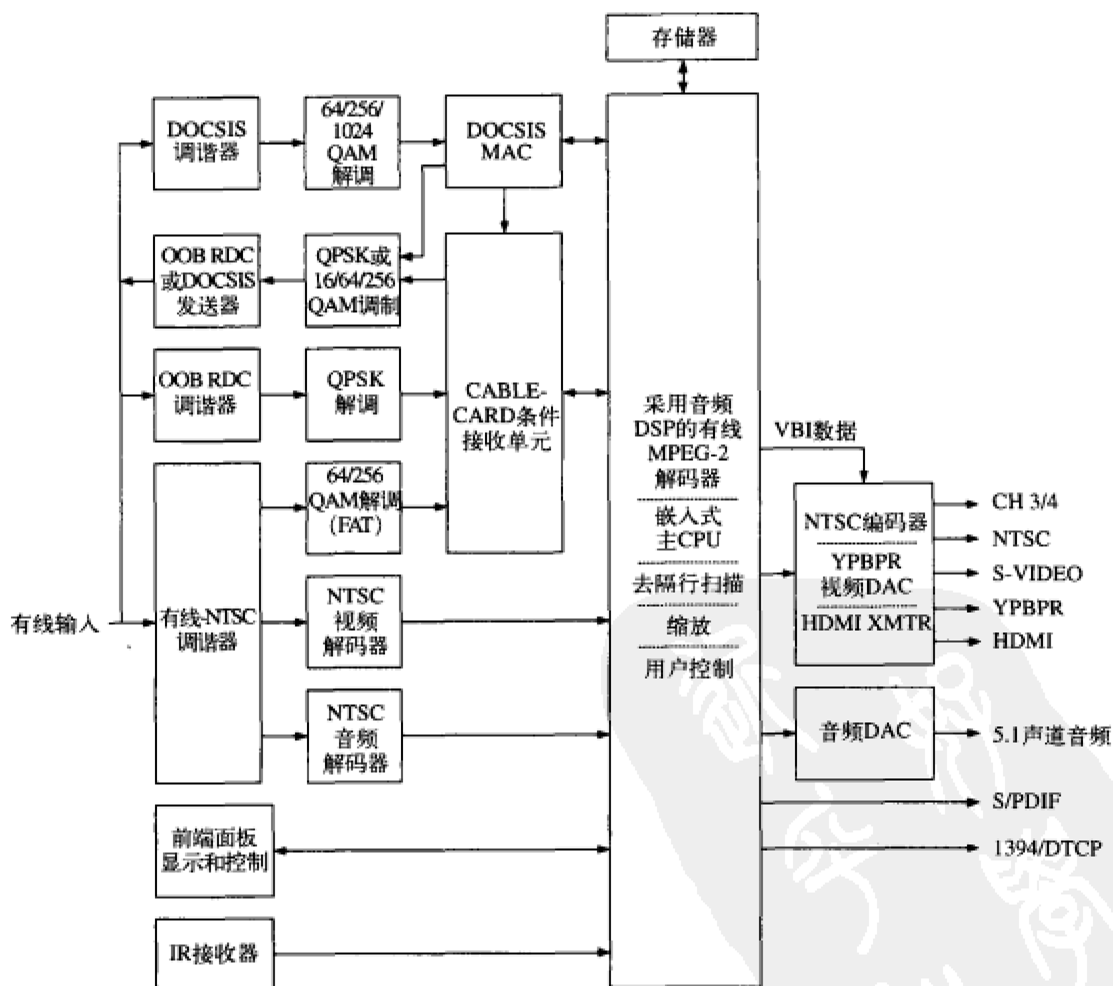
PacketCable™定义了一个发送实时多媒体服务的通用平台。基于DOCSIS®，PacketCable™采用因特网协议（IP）技术。

16.8　应用框图

图16-2描述了OpenCable™接收器机顶盒框图。机顶盒通常要求能够同时输出节目的高清和标清版本。

792

图16-3示出了包含一路OpenCable™接收器的数字电视的框图。



793

图16-2　典型的OpenCable™接收端机顶盒框图

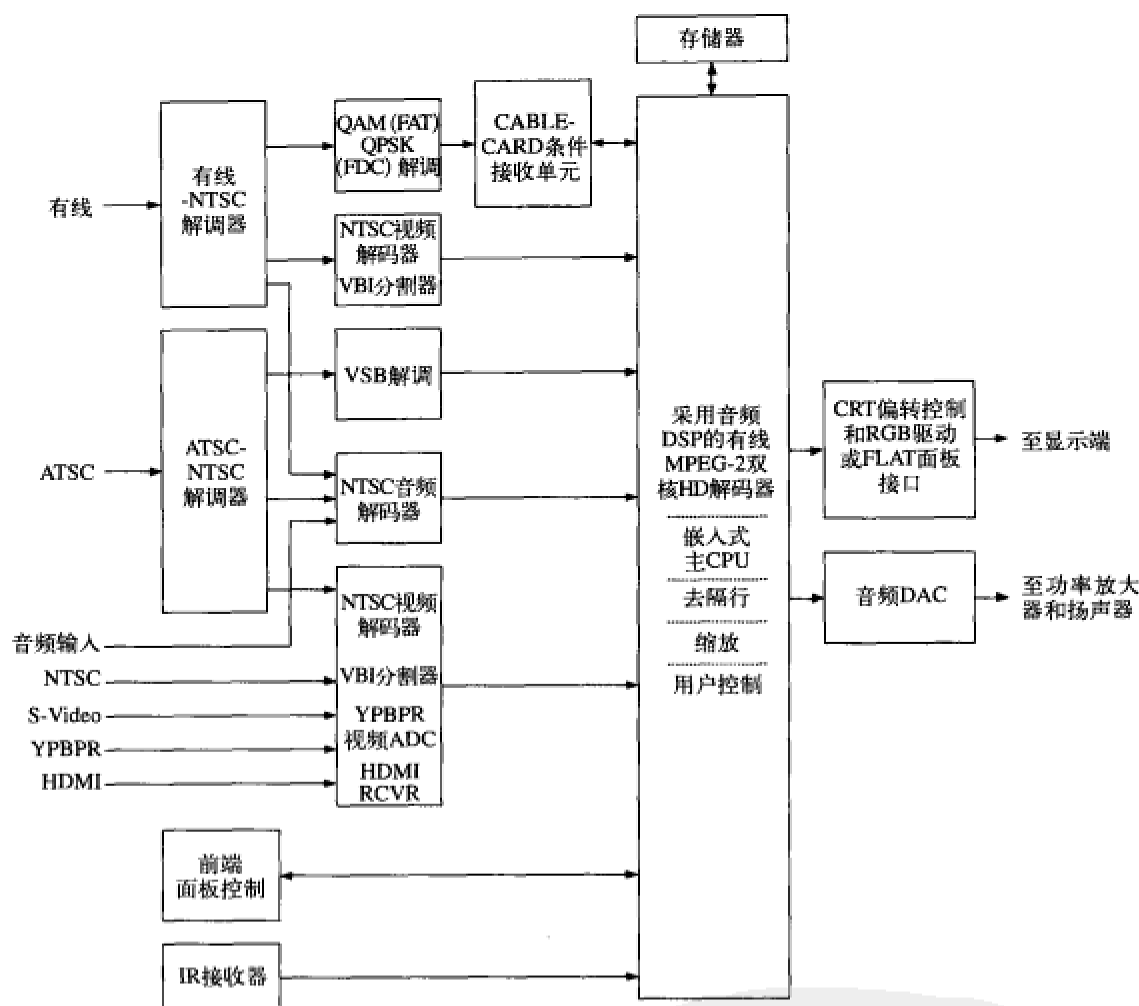


图16-3　数字电视中的OpenCable™接收器

794

参考文献

1. ATSC A/53E, *ATSC Digital Television Standard*, September 2006.
2. ATSC A/65C, *ATSC Standard: Program and System Information Protocol for Terrestrial Broadcast and Cable*, January 2006.
3. ATSC A/90, *ATSC Data Broadcast Standard*, July 2000.
4. SCTE 07, *Digital Transmission Standard for Cable Television*, 2006.
5. SCTE 28, *HOST-POD Interface Standard*, 2004.
6. SCTE 40, *Digital Cable Network Interface Standard*, 2004.
7. SCTE 54, *Digital Video Service Multiplex and Transport System for Cable Television*, 2004.
8. SCTE 65, *Service Information Delivered Out-of-Band for Digital Cable Television*, 2002.
9. SCTE 80, *In-Band Data Broadcast Standard Including Out-of-Band Announcements*, 2002.

795

第17章 DVB数字电视

除了美国、加拿大、韩国、阿根廷、巴西和中国台湾地区之外，大多数国家和地区都使用DVB（数字视频广播）数字电视（DTV）广播标准。

其他三个重要的DTV标准分别是ISDB（综合业务数字广播）、ATSC（先进电视系统委员会）和OpenCable™。三种标准的基本音频和视频能力非常相似。三种标准的主要区别在于RF调制机制和非音/视频服务的定义。表17-1对各种DVB标准进行了比较。

DVB标准实际上是一组ETSI标准。

EN 300 421——DVB-S，帧结构，针对11/12GHz卫星服务的信道编码和调制

EN 300 429——DVB-C，帧结构，有线系统的信道编码和调制

EN 300 468——定义DVB系统中的服务信息

EN 300 472——定义如何在DVB比特流中传输IRU-R系统B图文电视

EN 300 743——字幕系统

EN 300 744——DVB-T，帧结构，数字地面电视的信道编码和调制

EN 301 192——数据广播的DVB定义

EN 301 775——定义在DVB系统中如何传递垂直消影信息（VBI）数据

EN 302 304——DVB-H，针对手持终端的传输系统

EN 302 307——DVB-S2，第二代帧结构，针对广播、交互服务、新闻采集和其他宽带卫星服务的信道编码和调制系统

ES 200 800——针对有线电视分布系统（CATV）的DVB交互信道

ETS 300 801——通过公共交换电信网络（PSTN）/综合业务数字网（ISDN）实现的交互信道

ETS 300 802——针对DVB交互服务的、与具体网络无关的协议

TR 101 190——DVB地面服务及传输方面的实现方案

TR 101 194——用于DVB交互服务的、网络无关协议规范的实现与使用指南

TR 101 200——DVB规范和标准的使用方案

TR 101 202——数据广播的实现方案

TR 101 211——服务信息（SI）的实现和使用方案

TS 101 154——在基于MPEG-2传输流的广播应用中使用视频和音频编码的实现方案

TS 101 699——针对通用接口定义的扩展

TS 102 470——DVB-H上的数据广播，节目专用信息（PSI）/服务信息（SI）

TS 102 472——DVB-H上的数据广播，内容传输协议

EN 50221——针对条件接收和其他数字视频广播解码应用的通用接口定义

ETR 289——为了在数字广播系统中采用加扰和条件接收（CA）的相关支持

表17-1 各种数字电视标准的比较

参 数	DVB-T（地面）	DVB-C（有线）	DVB-S/-S2（卫星）	DVB-H（手持）	DVB-SH（手持）
视频压缩	MPEG-2、MPEG-4.10（H.264）、SMPTE 421M（VC-1）			MPEG-4.10（H.264）、SMPTE 421M（VC-1）	
音频压缩	MPEG、Dolby® Digital、Dolby® Digital Plus、DTS、MPEG-4 AAC、MPEG-4 HE-AAC v1/v2			MPEG-4 AAC、MPEG-4 HE-AAC v1/v2、AMR-WB+	

（续）

参　数	DVB-T（地面）	DVB-C（有线）	DVB-S/-S2（卫星）	DVB-H（手持）	DVB-SH（手持）
复用	MPEG-2传输流			进行RTP封装之后的MPEG-2传输流	
调制	COFDM	QAM	QPSK	采用DVB-T	采用DVB-S
信道带宽	6MHz、7MHz或8MHz	6MHz、7MHz或8MHz	–	6MHz、7MHz或8MHz	–

797

DVB使用MPEG-2传输流在6MHz、7MHz或8MHz信道上传输压缩的数字视频、数字音频和数据。MPEG-2传输流中可以同时存在多个视频流，多个音频流和数据。

MPEG-2传输流的最大比特流约为24.1Mbit/s（8MHz DVB-T）或约51Mbit/s（8MHz 256-QAM DVB-C）。DVB-S的比特率取决于收发机的带宽和采用的码率，最大值可达到54Mbit/s（DVB-S2的比特率在DVB-S的基础上可以提高25%~30%）。比特率的分配可以采用非常灵活的方式，具体需要在所提供的节目的数量和视频及分辨率之间进行权衡。

面向移动应用的DVB-H和DVB-SH分别采用DVB-T和DVB-S IP数据广播。传输流采用的是RTP打包之后的MPEG-2传输流。

第二代DVB-T和DVB-C（分别称为DVB-T2和DVB-C2）目前正在制定过程中。

17.1 视频能力

虽然，只要不超过最大比特率就可以采用任意分辨率，但通常只采用几个标准化的分辨率，如表17-2所示。对于大多数分辨率允许采用隔行或逐行两种方式。

DVB-T、DVB-C、DVB-S和DVB-S2支持MPEG-2（MP@ML、MP@HL）、MPEG-4.10（MP@L3、HP@L4）和SMPTE 421M（AP@L1、AP@L3）视频。

DVB-IP（“在IP上实现DVB”，DVB-H、DVB-SH和DVB-IPTV中采用）针对MPEG-4.10（BP@L1b、BP@L1.2、BP@L2）和SMPTE 421M（SP@L1、SP@ML、AP@L0）视频增加了额外的支持。

17.2 音频能力

DVB-T、DVB-C、DVB-S和DVB-S2支持MPEG-1层2、MPEG-2多通道层2、Dolby® Digital、Dolby® Digital Plus、DTS、MPEG-4 ACC和MPEG-4 HE-ACC v1/v2音频。

DVB-IP（“在IP上实现DVB”，DVB-H、DVB-SH和DVB-IPTV中采用）针对AMR-WB+增加了额外的支持。

17.3 系统信息

ETSI EN 300 468定义了系统信息（SI）数据，该数据称为DVB比特流的一部分。SI是一组用来扩展MPEG-2PSI表的层次化的相关表（见图17-1和表17-3）。它提供了其他传输流甚至其他网络中存在的数据的相关信息。对于用户而言，信息的表示方式没有定义，从而允许接收器制造商选择合适的表示方式。

17.3.1 所需表

1. 事件信息表（EIT）

事件信息表总共有128个（EIT0~EIT127），每个EIT描述每个信道上的事件或TV节目。每个

EIT的有效时间是3小时。由于总共有128个EIT，未来16天内的节目都可以提前预告。前4个EIT是必须的（建议的个数为24）。

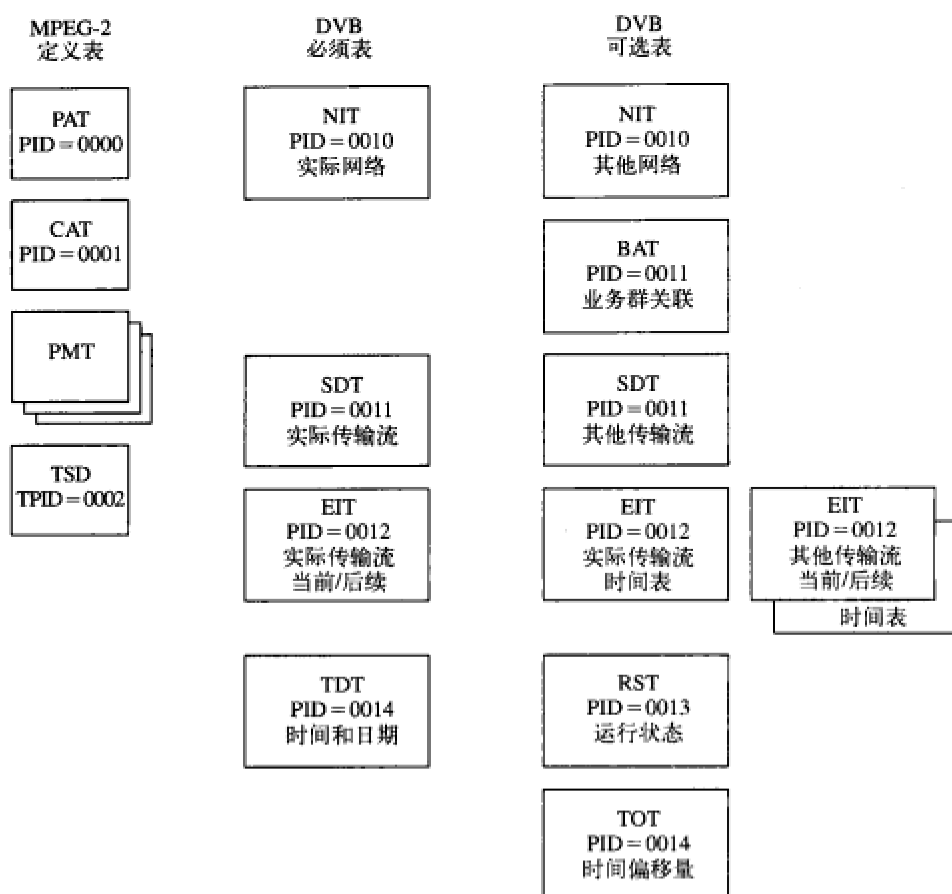


图17-1 DVB SI表的关系

EIT提供的信息包括开始时间、持续时间、标题、指向事件可选描述文本的指针、分级数据（advisory data）、字幕服务数据、音频服务描述符等。

798

2. 网络信息表（NIT）

网络信息表提供物理网络的相关信息，包括任何传输流分组和相关调制信息。它可以用于构建接收器，相关的调制信息存储在非易失性存储器中。NIT同样可以用于指示调制信息的变化。

3. 服务描述表（SDT）

SDT描述可用的服务，如服务名、服务提供商等。

4. 时间和日期表（TDT）

TDT包含实际的UTC时间，该时间被编码成修正的公元时间。接收器可以使用该信息来维持正确的本地时间。

17.3.2 可选表

1. 业务群关联表（BAT）

BAT提供了业务（可能穿过网络边界的服务的组合）的相关信息。除了业务名称，BAT还提

供了每个业务的服务列表。

2. 不连续信息表（DIT）

DIT出现在SI信息不连续时的转换点。该表的使用限制在局部传输流中，在广播中没有采用。

3. IP/MAC通告表（INT）

INT用来指示DVB网络中存在的IP流及其位置。INT表的数目可以不止一个。 799

表17-2 DVB数字电视中视频常见有效分辨率

有效分辨率（Y）	SDTV或HDTV	帧率（0＝逐行，1＝隔行）						
		23.976p 24p	25i	29.97i 30i	25p	29.97p 30p	50p	59.94p 60p
480×480	SDTV	×		×		×		×
480×576		×	×		×		×	
544×480		×		×		×		×
544×576		×	×		×		×	
704×480		×		×		×		×
704×576		×	×		×		×	
1280×720	HDTV	×			×	×	×	×
960×1080		×	×	×	×	×		
1280×1080		×	×	×	×	×		
1440×1080		×	×	×	×	×		
1920×1080		×	×	×	×	×		

表17-3 DVB SI表、描述符和描述符位置

描 述 符	描述符标识	表						
		PMT	NIT	BAT	SDT	EIT	TOT	SIT
PID		每PAT	0x0010	0x0011	0x0011	0x0012	0x0014	0x001F
table_ID		0x02	0x40 0x41	0x4A	0x42 0x46	0x4E~ 0x6F	0x73	0x7F
重复率		100ms	10s	10s	2~10s	2~10s	30s	30s
AAC	0111 1100	×						
AC-3	0110 1010	×						
适配域数据	0111 0000	×						
应用标识	0110 1111	×						
辅助数据	0110 1011	×						
声明支持	0110 1110				×			
业务名称	0100 0111			×	×			×
有线传输系统	0100 0100		×					
CA标识符	0101 0011			×	×	×		×
单元频率链接	0110 1101		×					
单元列表	0110 1100		×					
组件	0101 0000				×	×		×
内容	0101 0100					×		×
国家可用性	0100 1001			×	×			×
数据广播	0110 0100				×	×		×
数据广播ID	0110 0110	×						
DSNG	0110 1000							
DTS音频	0111 1011	×						
增强AC-3	0111 1010	×						

（续）

描　述　符	描述符标识	表						
		PMT	NIT	BAT	SDT	EIT	TOT	SIT
扩展事件	0100 1110					×		×
扩展	0111 1111	×	×	×	×	×	×	×
频率列表	0110 0010		×					
链接	0100 1010		×	×	×	×		×
本地时间偏移	0101 1000						×	
Mosaic	0101 0001	×			×			×
多语种业务名	0101 1100			×				
多语种组件	0101 1110					×		×
多语种网络名	0101 1011		×					
多语种服务名	0101 1101				×			×
网络名	0100 0000		×					
NVOD参考	0100 1011				×			×
家长分级控制	0101 0101					×		×
局部传输流	0110 0011							×
PDC	0110 1001					×		
私有数据说明符	0101 1111	×	×	×	×	×		×
卫星传输系统	0100 0011		×					
S2卫星传输系统	0111 1001		×					
加扰	0110 0101	×						
服务	0100 1000				×			×
服务可用性	0111 0010				×			
服务标识符	0111 0001				×			
服务列表	0100 0001		×	×				
服务切换	0110 0000	×						
短事件	0100 1101					×		×
短平滑缓冲	0110 0001					×		×
流标识符	0101 0010	×						
填充	0100 0010		×	×	×	×		×
字幕	0101 1001	×						
电话	0101 0111				×	×		×
图文电视	0101 0110	×						
地面传输系统	0101 1010		×					
时间平移事件	0100 1111					×		×
时间平移服务	0100 1100				×			×
传输流	0110 0111							
VBI数据	0100 0101	×						
VBI图文电视	0100 0110	×						

注：1. PMT：MPEG-2节目映射表。
　　2. SIT仅存在于局部传输流中。

800 ≀ 803

4. 运行状态表（RST）

RST更新一个或多个事件的运行状态。它们只在事件状态发生变化时向外发送一次，而其他的表通常重复传输。

5. 选择信息表（SIT）

SIT描述局部传输流中携带的服务和事件。该表的使用限制在局部传输流中。在广播中没有采用。

6. 填充表（ST）

ST用于替换、作废子表或使SI表完整。

7. 时间偏移表（TOT）

TOT类似于TDT，区别在于它包含了本地时间偏移信息。

17.3.3 描述符

和MPEG-2非常相似，DVB使用描述符来添加新功能。除了各种MPEG-2描述符之外，在PMT、一个或多个SI表（见表17-3）内可以包含一个或多个DVB专有描述符以扩展表内数据。解码器不能识别的描述符必须丢弃。这使得添加新功能不影响不能识别和处理描述符的接收端。

1. ACC描述符

该描述符指示提供了MPEG-4 ACC、HE-ACC、或HE-ACC v2音频，第13章对此进行了详细的讨论。

2. AC-3和增强AC-3描述符

这些DVB描述符指示音频格式为Dolby® Digital或Dolby® Digital Plus，第13章对此进行了详细的讨论。

3. 适配域数据描述符

这一描述符指示MPEG-2适配域的私有数据域内的数据域类型。第13章对此进行了详细的讨论。

4. 辅助数据描述符

这一描述符指示在MPEG音频基本流中是否存在附加数据及附加数据类型。第13章对此进行了详细的讨论。

5. 声明支持描述符

这一DVB描述符标识服务支持的声明的类型。它同时还标识声明传输方式并提供链接信息，使得可以对声明流进行监测。

6. 业务名称描述符

该DVB描述符以可变长文本的形式提供了业务名称，如Max Movie Channels。

7. CA标识符描述符

该DVB描述符标识是否有一个业务、服务或事件和条件接收系统相关联，如果有，标识出所使用的条件接收。

8. 有线传输系统描述符

该DVB描述符用于传输有线网络的物理参数，包括频率、调制方式和符率。

804

9. 单元频率链接描述符

该DVB描述符用于NIT中以描述地面网络。它为传输流提供了单元与该单元使用的频率之间的链接。

10. 单元列表描述符

该DVB描述符提供了一个所有网络单元的列表，NIT指示并描述了这些单元的覆盖面积。

11. 组件描述符

该DVB描述符指示流的类型，它可以用于提供对流的文本描述。第13章对此进行了详细的讨论。

12. 内容描述符

该DVB描述符用于确认内容的类型（喜剧、现场访谈等）。

13. 国家可用性描述符

该DVB描述符标识出允许或不允许接收该服务的国家。对于每个服务，该描述符可以出现两次，第一次列出允许接收该服务的国家，第二次列出不允许接收该服务的国家。后续的列表将覆盖前面的列表。第13章对此进行了详细的讨论。

14. 数据广播描述符

该DVB描述符标识出SI内可用的数据广播服务。

15. 数据广播ID描述符

该DVB描述符指示数据编码系统标准。它是数据广播描述符的精简形式，并且可以存在于PMT中。第13章对此进行了详细的讨论。

16. DSNG描述符

该DVB描述符只存在于DSNG（数字卫星新闻采集）传输中。

17. DTS音频描述符

该DVB描述符表示存在DTS音频基本流。第13章对此进行了详细的讨论。

18. 扩展事件描述符

该DVB描述符提供了对事件的文本描述，它可以和短事件描述符一起使用。为了阐述多于256字节的信息，可以采用多个描述符。

19. 扩展描述符

该DVB描述符扩展了描述符标志域的8位的值。第13章对此进行了详细的讨论。

20. 频率列表描述符

该DVB描述符可以存在于NIT中。当内容采用其他的频率传输时，它提供了额外的频率。

21. 链接描述符

该DVB描述符提供了一个指向其他服务、传输流、节目导播、服务信息、软件更新等的链接。

22. 本地时间偏移描述符

该DVB描述符可以存在于TOT中，用以描述与特定国家相关的相对于UTC的本地时间偏移的动态变换。它使得接收器可以自动在夏季和冬季时间之间进行调整。

805

23. Mosaic描述符

该DVB描述符将数字视频分割成基本单元，并控制基本单元向逻辑单元的分配，提供逻辑单元的内容和对应信息之间的链接（例如，业务、服务、事件等）。第13章对此进行了详细的讨论。

24. 多语种业务名称描述符

该DVB描述符以一种或多种语言提供了业务名称的文本形式。

25. 多语种组件描述符

该DVB描述符以一种或多种语言提供了组件名称的文本形式。各组件通过组件标识值识别。

26. 多语种网络名称描述符

该DVB描述符以一种或多种语言提供了网络名称的文本形式。

27. 多语种服务名称描述符

该DVB描述符提供了服务提供商的名称，并以一种或多种语言提供了服务的文本形式。

28. 网络名描述符

该DVB描述符以文本形式传输网络名。

29. NVOD（准视频点播）参考描述符

该描述符，协同时间平移服务描述符和时间平移事件描述符，提供了一种有效地描述服务的

方式，这些服务携带相同的事件序列，具有不同的起始时间偏移。

30. 家长分级控制（Parental Rating）描述符

该DVB描述符提供了一种基于年龄的分类方式，并且提供了相关扩展以使得可以采用其他分类标准。第13章对此进行了详细的讨论。

31. 局部传输流描述符

SIT包含了用于控制、播放、复制局部传输流的所有信息。该描述符描述了这些信息。

32. PDC描述符

该DVB描述符扩展了DVB系统，使其增加了PDC（节目传输控制）功能。PDC由ETSI EN 300 231和ITU-R BT.809定义，第8章对此进行了详细讨论。

33. 私有数据说明符描述符

该DVB描述符用于标识任何私有描述符或描述符中的私有域的来源。第13章对此进行了详细的讨论。

34. 卫星传输系统描述符

该DVB描述符用于传输卫星网络的物理参数，包括频率、轨道位置、东西标志、极性、调制模式和符率。

35. 加扰描述符

该DVB描述符指示加扰系统所选择的操作模式。第13章对此进行了详细的讨论。

36. 服务描述符

该DVB描述符以文本形式提供了服务和服务提供商的名称。 806

37. 服务可用性描述符

该DVB描述符存在于地面网络的SDT中。它表示服务对识别出的单元是否可用。

38. 服务列表描述符

该DVB描述符对每个传输流提供了一个服务和服务类型列表。

39. 服务切换描述符

当服务从一个传输流切换到另一个传输流时，该DVB描述符为接收器提供了一种跟随服务的方式。这种切换过程将会对视频和音频产生一些干扰。第13章对此进行了详细讨论。

40. 短事件描述符

该DVB描述符提供了事件名称以及对事件的简短描述。

41. 短平滑缓冲描述符

该MPEG-2描述符使得可以在PSI中表示出服务的比特率。

42. 流标识描述符

该DVB描述符使得流可以和EIT中的一个描述相关联。当一个服务中有多个流具有相同的类型时，该描述符非常有用。第13章对此进行了详细讨论。

43. 填充描述符

该DVB描述符用于填充表格的内容。进行填充的原因可以是各种各样的。它也可以用于禁用已经失效的描述符。

44. 字幕描述符

该DVB描述符用于识别ETSI EN 300 743字幕数据。第13章对此进行了详细讨论。

45. 电话描述符

该DVB描述符给出了一个电话号码，它可以和一个PSTN或有线调制解调器一起支持窄带交互信道。

46. 图文电视描述符

该DVB描述符用于标识携带EBU图文电视数据的基本流。第13章对此进行了详细讨论。

47. 地面传输系统描述符

该DVB描述符用于传输地面网络的物理参数，包括中心频率、带宽、卫星、层次结构、码率、防护间隔和传输模式。

48. 时间平移事件描述符

该DVB描述符表示一个事件是另一个事件的时间平移后的副本。

49. 时间平移服务描述符

该DVB描述符将一个服务链接到最多20个携带相同节目的服务，原节目和链接的节目在时间上有一个平移。典型的应用是类似视频点播（NVOD）服务。

50. 传输流描述符

该DVB描述符只在TSDT中传输，用以表示MPEG传输流与DVB或DSNG相兼容。

51. VBI数据描述符

807

该DVB描述符定义了在相关基本流中的VBI服务类型。第13章对此进行了详细讨论。

52. VBI图文电视描述符

该描述符的语法类似于*图文电视描述符*，区别在于它没有用于将stream_type＝0x06与VBI或EBU图文电视标准相关联。解码器采用该描述符中的语言选择杂志（magazine）和字幕。

17.4　数据广播

DVB数据广播标准描述了在DVB流中传输数据时可用的封装协议。它基于MPEG-2 DSM-CC定义了DVB数据管道、DVB数据流、DVB多协议封装、DVB数据轮播和DVB目标轮播。DVB增加了特定的信息，以使得DSM-CC框架可以在DVB环境尤其是DVBSI中工作。

目前，数据广播有5种不同的应用领域，每种对广播都有不同的要求。对于每个应用领域，定义了一个档次和附加的描述符用于支持该应用领域。第13章中包含了附加的数据广播信息。

应用领域和档次

1. 数据管道

该档次支持采用简单、异步或点对点数据传输的数据广播服务。数据直接在MPEG-2传输流的载荷中传输。

2. 数据流

该档次支持采用面向流的或点对点数据传输的数据广播服务，数据传输方式可以是异步或同步方式。数据由MPEG-2 PES包携带。

3. 多协议封装

该档次支持采用某种通信协议的报文传输的数据广播服务，通信协议可以是IP多点传送。数据报文的传输是通过将报文封装在MPEG-2 DSM-CC段中实现的。

4. 数据轮播

该档次支持的数据广播服务采用了周期传输数据块的形式。这些块的长度固定，并且可以及时地从数据轮播中进行更新、添加或删除。数据采用MPEG-2 DSM-CC数据轮播的形式广播。

5. 目标轮播

该档次支持的数据广播服务采用了周期广播用户–用户（U-U）目标的形式。数据采用MPEG-2 DSM-CC目标轮播和DSM-CC数据轮播的形式广播。

17.5 条件接收

条件接收（CA）是指在传输之前对内容进行加密，以使得只有授权用户才能享受服务。为了对受保护的内容进行解密，接收器采用了CA模块。CA模块使得只有被授权的节目才能解密。

实现DVB条件接收有两种基本方式：同密和多密。

808

17.5.1 同密

同密技术取决于DVB通用加扰算法（CSA）。CSA是一种对传输流或程序基本流进行安全加扰和解密的工具。CSA解密电路集成在视频解码芯片中，而不是在CA模块内。CA模块并不直接对加密数据进行压缩，它仅仅向CSA解密电路提供解密的关键信息，使得流可以正确解密。

由于基于同密技术的接收器不需要使用DVB通用接口，可以将CA模块嵌入在接收器中，也可以基于NRSS-A接口采用可分离的CA模块。由于访问嵌入式CA模块的难度更大，嵌入式CA模块提高了安全性。然而，接收器难以进行更新或调整以应用于其他CA系统中，因此这种CA模块也可能会迅速成为系统的障碍。

采用基于EIA-679 NRSS-A接口（ISO 7816）的可分离CA模块使得修改CA系统就像更改低成本的CA卡一样容易。对于每个加扰器，CA卡由服务提供商提供。

同密技术还允许在一个广播或接收器中使用多个CA系统。每个CA系统的ECM（Entitlement Control Messages，权利控制消息）和EMM（Entitlement Management Messages，权利管理消息）在流中传输。接收器识别和使用解密所需要的ECM和EMM。因此，包含针对多个CA系统的数据的广播可以在接收器上查看，接收器只需要支持这些CA系统的任何一个。同密技术还使得开发新的CA解决方案的同时可以保持与旧CA系统的兼容性。

第二代DVB-CSA（称为DVB-CSA2）正在制定过程中。

17.5.2 多密

多密技术是一个基于可分离CA模块的开放系统，它由每个加扰器的服务供应商提供。加密后的流发送到CA模块。CA模块直接从流中找到并抽取出所需数据，如ECM、EMM。解密后的流输送到MPEG-2解码器。

CA模块通过DVB通用接口插入到接收器中。多密技术的优点在于接收器可以方便地进行配置，以采用不同的、非兼容的CA系统接收来自不同的服务提供商的服务。因此，接收器不会成为障碍。

通过采用多个DVB通用接口，接收器可以支持多个CA系统，以支持多个CA模块。加密的数据顺序地通过不同的CA模块，每个CA模块从流中抽取出各自的ECM和EMM。

17.5.3 DVB通用接口

DVB通用接口是接收器和CA功能模块的接口，它分离了接收器和CA模块。它是基于EIA-679 NRSS-B接口（PCMCIA或PC卡）的，是多密系统的关键。

传输流接口由8位并行输入、8位并行输出、控制信号和时钟构成。命令（主机）接口由8位双向数据总线、地址和控制信号组成。

DVB通用接口同样可以用来向接收器增加新的特性，如支持一种新的音频编解码器或增加视觉消影音频能力。

第二代DVB-CI（称为DVB-CI2）正在制定过程中。

809

17.6 应用框图

图17-2和图17-3示出了一个典型的DVB-S机顶盒结构框图。

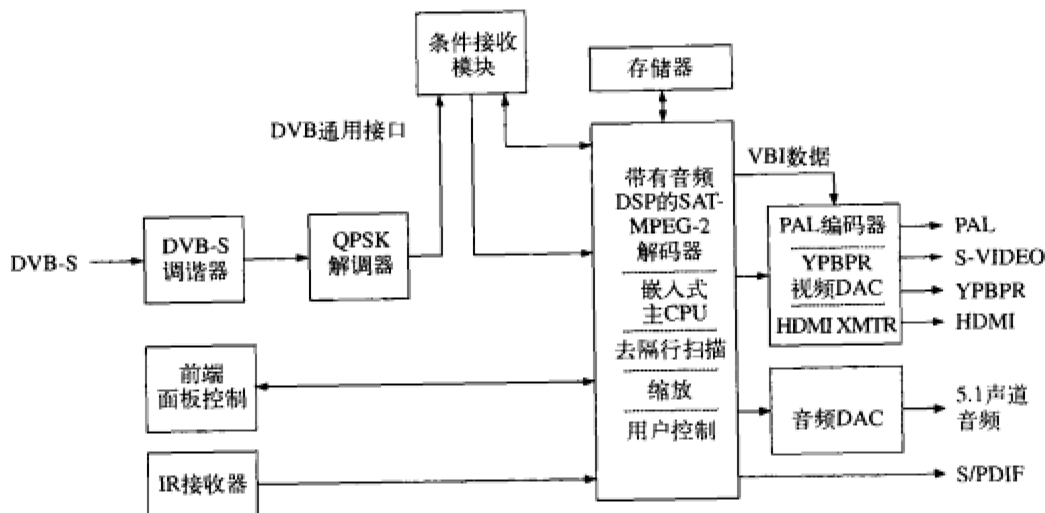


图17-2 DVB接收器机顶盒框图（多密）

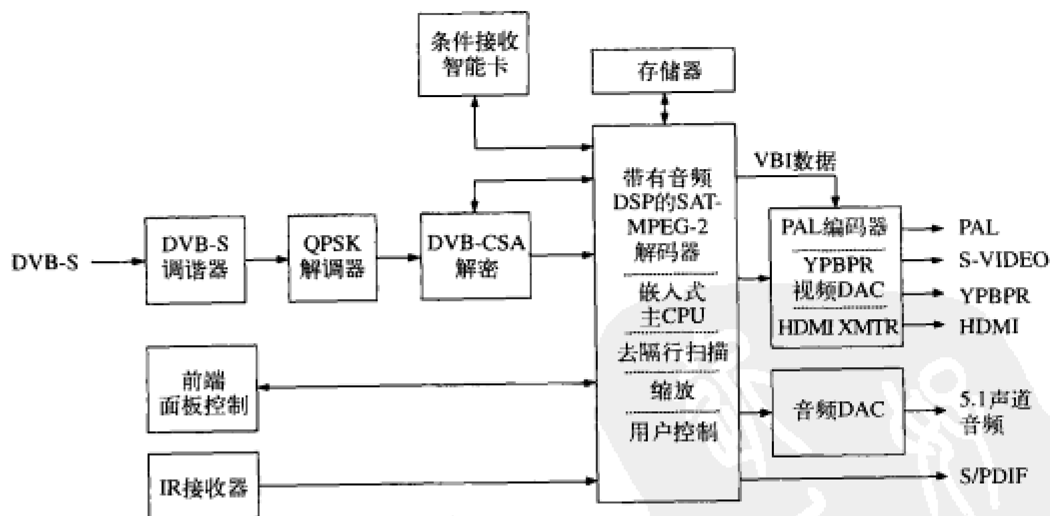


图17-3 DVB接收器机顶盒框图（同密）

参考文献

1. ETSI EN 300 421, *Digital Video Broadcasting (DVB): Framing Structure, Channel Coding and Modulation for 11/12 GHz Satellite Services*, August 1997.
2. ETSI EN 300 429, *Digital Video Broadcasting (DVB): Framing Structure, Channel Coding and Modulation for Cable Systems*, April 1998.
3. ETSI EN 300 468, *Digital Video Broadcasting (DVB): Specification for Service Information (SI) in DVB Systems*, May 2006.
4. ETSI EN 300 472, *Digital Video Broadcasting (DVB): Specification for Conveying ITU-R System B Teletext in DVB Bitstreams*, May 2003.
5. ETSI EN 300 743, *Digital Video Broadcasting (DVB): Subtitling Systems*, November 2006.

6. ETSI EN 300 744, *Digital Video Broadcasting (DVB): Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*, November 2004.
7. ETSI EN 301 192, *Digital Video Broadcasting (DVB): DVB Specification for Data Broadcasting*, November 2004.
8. ETSI EN 301 775, *Digital Video Broadcasting (DVB): Specification for the Carriage of Vertical Blanking Information (VBI) Data in DVB Bitstreams*, May 2003.
9. ETSI ES 200 800, *Digital Video Broadcasting (DVB): DVB Interaction Channel for Cable TV Distribution Systems (CATV)*, October 2001.
10. ETSI ETS 300 801, *Digital Video Broadcasting (DVB): Interaction Channel through Public Switched Telecommunications Network (PSTN) / Integrated Services Digital Networks (ISDN)*, August 1997.
11. ETSI ETS 300 802, *Digital Video Broadcasting (DVB): Network-Independent Protocols for DVB Interactive Services*, November 1997.
12. ETSI TR 101 190, *Digital Video Broadcasting (DVB): Implementation Guidelines for DVB Terrestrial Services; Transmission Aspects*, November 2004.
13. ETSI TR 101 194, *Digital Video Broadcasting (DVB): Guidelines for Implementation and Usage of the Specification of Network Independent Protocols for DVB Interactive Services*, June 1997.
14. ETSI TR 101 200, *Digital Video Broadcasting (DVB): A Guideline for the Use of DVB Specifications and Standards*, September 1997.
15. ETSI TR 101 202, *Digital Video Broadcasting (DVB): Implementation Guidelines for Data Broadcasting*, January 2003..
16. ETSI TR 101 211, *Digital Video Broadcasting (DVB): Guidelines on Implementation and Usage of Service Information (SI)*, February 2006.
17. ETSI TS 101 154, *Digital Video Broadcasting (DVB): Implementation Guidelines for the Use of Video and Audio Coding in Broadcasting Applications Based on the MPEG-2 Transport Stream*, June 2005.
18. ETSI TS 101 699, *Digital Video Broadcasting (DVB): Extensions to the Common Interface Specification*, November 1999.

810 ~ 811



第18章　ISDB数字电视

ISDB（集成服务数字广播）数字电视（DTV）广播标准用于日本。

其他三个主要的DTV标准是ATSC（先进电视系统委员会）、DVB（数字视频广播）和OpenCable™。三种标准的基本音频和视频能力非常相似。三种标准的主要区别在于RF调制机制和非音/视频服务的定义。ISDB基于DVB，增加了日本要求的额外服务。各种ISDB标准的比较见表18-1。

表18-1　ISDB标准之间的比较

参　数	ISDB-T（地面）	ISDB -C（有线）	ISDB-S（卫星）
视频压缩	MPEG-2，MPEG-4.10（H.264）		
音频压缩	MPEG-2 AAC，MPEG-4 HE-AAC		
复用	MPEG-2传输流		
调制	BST-OFDM*	QAM	PSK
信道带宽	6MHz、7MHz或8MHz	6MHz、7MHz或8MHz	—

注：BST-OFDM = OFDM 频带分段传输。

ISDB标准实际上是一组ARIB标准：

STD-B10——针对数字广播系统的服务信息

STD-B16——标准数字接收器，通常用于采用卫星通信的数字卫星广播服务

STD-B20——ISDB-S，数字卫星广播的传输系统

STD-B21——数字广播的接收器（理想定义）

STD-B23——针对数字广播的应用执行引擎平台

STD-B24——针对数字广播的数据编码和传输定义

STD-B25——针对数字广播的条件接收系统定义

STD-B31——ISDB-T，针对数字地面电视广播的传输系统

STD-B32——针对数字广播的视频编码、音频编码和复用定义

STD-B40——针对附加数据的PES包传输机制

ISDB采用MPEG-2传输流传输压缩视频、压缩音频和数据。类似于DVB，传输流可以通过地面、有线或卫星进行传播。交互式应用是基于BML（Broadcast Mark-up Language）的。

18.1　ISDB-S（卫星）

卫星标准有两个：ISDB-S［又称为BS（broadcast satellite，广播卫星）系统］和DVB-S［又称为CS（communication satellite，通信卫星）系统］。

ITU-R BO.1408也定义了ISDB-S（BS）。采用TC8PSK调制和34.5MHz的收发器时，最大比特率约为52.2Mbit/s。

CS每个传输通道只支持一个传输流，在27MHz的信道上支持的比特率可达到34Mbit/s。类似于DVB-S，调制策略采用的是QPSK。不同于ISDB的其他版本的是，CS采用的是MPEG-2 MP@ML视频（480i或480p）和MPEG-2 BC音频。

18.2 ISDB-C（有线）

ISDB-C采用64-QAM调制，有两个版本，其中一个版本在一个传输信道上只支持一个传输流，另一个版本在一个传输信道上可以支持多个传输流。在一个6MHz的信道上，ISDB-C可以传输的比特率可达29.16Mbit/s。由于ISDB-S卫星信道上的比特率是ISDB-C的两倍，采用两个有线信道可以传输卫星信息。

由于信道带宽是一样的，所以ISDB-C同样允许传输基于OFDM的ISDB-T信号。

ISDB-S（BS）信号可以通过在数据转发器端对卫星信号进行分解，然后在接收端进行合成的方式传输。这种技术只适合于拥有很多空闲信道（可达29个）的有线系统。

18.3 ISDB-T（地面）

ISDB-T，地面广播标准，也是由ITU-R BT.1306定义的。采用5.6MHz的带宽，最大比特率可以达到23.2Mbit/s。ISDB-T同样支持6MHz、7MHz、8MHz的带宽。

带宽被分为13个OFDM分段，每个分段可以归入3个不同的分段组（分级层），每个组有不同的传输参数，如载波调制策略、内部码编码率、时间间隔长度等。这使得采用不同分辨率进行传输的同样的节目，可以采用固定接收器显示高清图像的同时采用移动接收器显示标清图像。

18.4 视频能力

表18-2列出了几个标准分辨率。

早期的视频压缩是基于MPEG-2 MP@ML或MP@HL的。然而，对于MPEG-2的参数有一些约束，这一点在MPEG-2这一章中进行了详细讨论。

分辨率为176×144（64 kbit/s或384 kbit/s）或325×288（128 kbit/s、384 kbit/s或2000 kbit/s）的MPEG-4.2简单档次或核心档次视频同样也支持。

分辨率为176×144（64 kbit/s）或325×288（192 kbit/s、384 kbit/s、2000 kbit/s或4000 kbit/s）的MPEG-4.10=（H.264）基准档次或主要档次视频格式也支持。

表18-2 ISDB数字电视中常见有效视频分辨率

有效分辨率（Y）	SDTV或HDTV	帧率（P=逐行，I=隔行）				MPEG-2	MPEG-4.2	MPEG-4.10（H.264）
		23.976p 24p	29.97i 30i	29.97p 30p	59.94p 60p			
176×120	SDTV	×		×	×	×		
176×144		×		×	×	×	×	×
352×240		×		×	×	×		
352×288		×		×	×	×	×	×
352×480			×			×		
480×480			×			×		
544×480			×			×		
720×480			×			×		
1280×720	HDTV	×		×	×	×		
1440×1080		×	×	×		×		
1920×1080		×	×	×		×		

18.5 音频能力

早期的音频压缩是采用MPEG-2 AAC-LC实现的，最大可达到5.1声道。ISDB同样支持MPEG-4 HE-AAC音频。

18.6 静态图像能力

静态图像采用JPEG（ISO/IEC 10918-1）、PNG（Portable Network Graphics）、MNG（Multiple-image Network Graphics）、MPEG-2 I-帧、MPEG-4.2 I-VOP和MPEG-4.10（H.264）I-图像格式。

18.7 图形能力

图形命令包括Domain、Texture（填充、垂直方格、水平方格和交叉方格）、set color（前端、后端）、select color、Blink、Set Pattern、Point、Line（实线、虚线、间断线、虚间断线）、Arc（轮廓、填充）、Rectangle（轮廓、填充）和Polygon（轮廓、填充）。

814 图18-1示出了ISDB采用的5平面视频/图像结构。

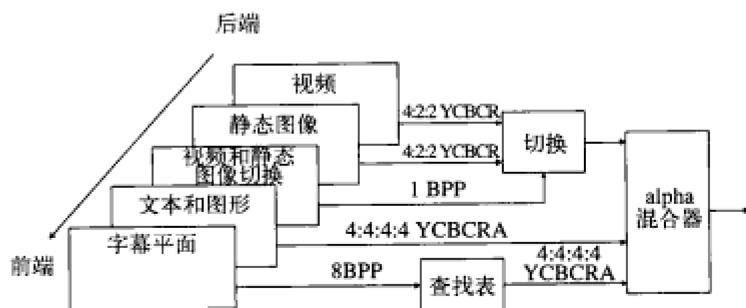


图18-1 ISDB 5平面视频/图形标准

18.8 系统信息

ARIB STD-B10定义了系统信息（SI）数据，该数据成为ISDB比特流的一部分。SI是一组用来扩展MPEG-2 PSI表的层次化的相关表（见表18-3）。它提供了其他传输流甚至其他网络中存在的数据的相关信息。对于用户而言，信息的表示方式没有定义，从而允许接收器制造商选择合适的表示方式。

18.8.1 表

1. 应用信息表（AIT）

AIT传输关于ARIB-J应用和执行情况的动态控制信息。

2. 业务群关联表（BAT）

BAT提供了业务（可能穿过网络边界的服务的组合）的相关信息。除了业务名称，BAT还提供了每个业务的服务列表。

3. 广播商信息表（BIT）

BIT用于在网络上提交广播商信息。

4. 通用数据表（CDT）

CDT传输所有接收端所需要的信息，这些信息将存储在非易失性存储器中。

5. 不连续信息表（DIT）

DIT出现在SI信息不连续时的转换点。该表的使用限制在局部传输流中，在广播中没有采用。

6. 下载表（DLT）

DLT传输用于下载的软件。

7. 下载控制表（DCT）

DCT传输的信息用于表示如何处理DLT。

8. 事件信息表（EIT）

事件信息表总共有128个，包括EIT0 ~ EIT127，每个EIT描述每个信道上的事件或TV节目。每个EIT的有效时间是3小时。由于总共有128个EIT，所以未来16天内的节目都可以提前预告。前4个EIT是必需的（建议使用前24个）。

EIT提供的信息包括开始时间、持续时间、标题、指向事件可选描述文本的指针、咨询数据、字幕服务数据、音频服务描述符等。

9. 事件关系表（ERT）

ERT用于表示节目或事件和它们的特性之间的关系。

10. 索引传输表（ITT）

ITT用于传输节目的节目索引信息。

11. 链接描述表（LDT）

LDT用于链接来自其他表的各种描述。

12. 局部事件信息表（LIT）

LIT传输节目中与局部事件相关的信息。局部事件包括时间、名称以及对事件的解说等。

815 ~ 816

13. 网络平台信息表（NBIT）

NBIT（Network Board Information Table）传输网络上的平台信息，如引导信息。

14. 网络信息表（NIT）

NIT提供物理网络的相关信息，包括传输流的分组信息和相关的调谐信息。它可以用于接收器启动阶段，相关的调谐信息存储在非易失性存储器中。NIT也可以用于标识调谐信息的变化。

15. 局部内容声明表（PCAT）

PCAT数据广播的局部内容声明。

16. 运行状态表（RST）

RST更新一个或多个事件的运行状态。它们只在事件状态发生变化时向外发送一次，而其他的表通常重复传输。

17. 选择信息表（SIT）

SIT描述局部传输流中携带的服务和事件。该表的使用限制在局部传输流中，在广播中没有采用。

18. 服务描述表（SDT）

SDT描述与信道相关的信息，如信道名、广播公司名称等。

19. 软件下载触发器表（SDTT）

SDTT传输标注信息，如下载服务ID、调度信息和用于校正的接收器类型。

20. 填充表（ST）

ST用于替换或作废子表，或使SI表完整。

21. 时间和日期表（TDT）

TDT包含实际的UTC时间，该时间被编码成修正的公元时间。接收器可以使用该信息来维持

正确的本地时间。

22. 时间偏移表（TOT）

TOT类似于TDT，区别在于它包含了局部时间偏移信息。

18.8.2　描述符

和MPEG-2非常相似，ISDB采用描述符来添加新功能。除了各种MPEG-2描述符之外，在PMT或一个或多个SI表（见表18-3）内可以包含一个或多个ISDB专有描述符，以扩展表内数据。解码器不能识别的描述符必须丢弃。这使得添加新功能不影响不能识别和处理描述符的接收端。

1. 音频组件描述符

该ARIB描述符描述一个音频基本流的参数。

2. AVC时序和HRD描述符

该描述符，如第13章所述，描述了视频流的时序信息和H.264（MPEG-4.10）的参考解码器。

3. AVC视频描述符

817

该ARIB描述符，如第13章所述，描述了H.264（MPEG-4.10）视频流的基本编码参数。

表18-3　ISDB SI表、描述符和描述符位置

描　述　符	描述符标识	表								
		PMT	NIT	BAT	SDT	EIT	TOT	BIT	NBIT	LDT
PID		每PAT	0x0010	0x0011	0x0011	0x0012	0x0014	0x0024	0x0025	0x0025
Table_ID		0x02	0x40 0x41	0x4A	0x42 0x46	0x4E~ 0x6F	0x73	0xC4	0xC5 0xC6	0xC7
重复率		100ms	10s	10s	2~10s	2~10s	30s	20s	20s	20s
音频组件	1100 0100				×					
AVC时序和HRD	0010 1010	×								
AVC视频	0010 1000	×								
基本局部事件	1101 0000									
平台信息	1101 1011								×	
业务名称	0100 0111			×	×					
广播商名称	1101 1000							×		
CA协约信息	1100 1011									
CA EMM TS	1100 1010									
CA标志符	0101 0011			×	×	×				
CA服务	1100 1100									
有线TS分配系统	1111 1001									
有线分布系统	0100 0100									
轮转兼容组合描述符	1111 0111	×				×				
组件	0101 0000	×				×				
组件组	1101 1001					×				
条件重放	1111 1000	×								
有连接传输	1101 1101									
内容	0101 0100					×				
内容可用性	1101 1110	×		×	×					
国家可用性	0100 1001	×		×	×					
数据组件	1111 1101	×								
数据内容	1100 0111					×				
数字复制控制	1100 0001	×			×	×				

（续）

描述符	描述符标识	表								
		PMT	NIT	BAT	SDT	EIT	TOT	BIT	NBIT	LDT
下载内容	1100 1001									
紧急信息	1111 1100	×	×			×				
事件组	1101 0110					×				
扩展广播商	1100 1110							×		
扩展事件	0100 1110					×				×
层次化传输	1100 0000	×								
超链接	1100 0101					×				
LDT链接	1101 1100					×				
链接	0100 1010	×	×	×	×	×				
本地时间偏移	0101 1000						×			
logo传输	1100 1111				×					
Mosaic	0101 0001	×			×					
网络标识符	1100 0010									
网络名称	0100 0000		×							
节点关系	1101 0010									
NVOD参考	0100 1011				×					
家长分级控制	0101 0101	×				×				
部分接收	1111 1011		×							
部分传输流	0110 0011									
参考	1100 0011									
卫星传输系统	0100 0011		×							
序列	1101 0101					×				
服务	0100 1000				×					
服务列表	0100 0001		×	×						
短事件	0100 1101					×				×
短节点信息	1101 0011									×
SI参数	1101 0111							×		
SI Prime_TS	1101 1010							×		
STC参考	1101 0100									
流标志符	0101 0010	×								
填充信息	0100 0010		×	×	×	×			×	×
系统管理	1111 1110	×	×							
目标区域	1100 0110	×								
地面传输系统	1111 1010		×							
时间－平移事件	0100 1111					×				
时间－平移服务	0100 1100				×					
TS信息	1100 1101									
视频解码控制	1100 1000	×								

注：PMT：MPEG-2节目映射表。

818～820

4. 基本局部事件描述符

该ARIB描述符表示局部事件标志符信息。

5. 平台信息描述符

该ARIB描述符以文本格式描述平台信息的标题和内容。

6. 业务名称描述符

该ARIB描述符以可变长文本的形式提供了业务名称，如Max Movie Channels。DVB也采用了

该描述符。

7. 广播商名称描述符

该ARIB描述符指示广播商的名称。

8. CA协约信息描述符

该ARIB描述符描述了针对设定节目的条件接收服务类型。

9. CA EMM TS描述符

该ARIB描述符描述了当EMM传输是通过特定的音井方法进行时特定的音井。

10. CA标志符描述符

该ARIB描述符标识是否有一个业务、服务或事件和条件接收系统相关联，如果有，标志出所采用的条件接收。

11. CA服务描述符

该ARIB描述符传输了提供自动消息指示服务的广播服务提供商的信息。

12. 轮转兼容组合描述符

该ARIB描述符，第13章中也有讨论，将在数据轮转传输定义（ARIB STD-B24第三部分）中定义的描述符作为子描述符，并且应用子描述符的功能描述了累加控制。

13. 组件描述符

该ARIB描述符（第13章中也有讨论）表示流的类型，它可以用于对流进行文本描述。DVB也采用了该描述符。

14. 组件组描述符

该ARIB描述符定义和标识一个事件中的组件组。

15. 条件播放描述符

该ARIB描述符（第13章中也有讨论）传递条件播放的描述和传输ECM与EMM的PID。

16. 有连接传输描述符

该ARIB描述符表示当接入地面音频传输路径中的传输时的物理条件。

17. 内容描述符

该ARIB描述符用于确认内容的类型（喜剧、现场访谈等）。DVB也采用了该描述符。

18. 内容可用性描述符

该ARIB描述符描述了控制接收器记录和输出内容的信息。encryption_mode标志表示是否对数字视频输出进行加密。它和数字复制控制描述符一起使用。

821

19. 国家可用性描述符

该ARIB描述符标识出允许或不允许接收该服务的国家。对于每个服务，该描述符可以出现两次，第一次列出允许接收该服务的国家，第二次列出不允许接收该服务的国家。后续的列表将覆盖前面的列表。第13章对此进行了详细讨论。

20. 数据组件描述符

该ARIB描述符标识出数据组件。第13章对此进行了详细讨论。

21. 数据内容描述符

该ARIB描述符描述了数据广播事件中每个内容相关的信息。

22. 数字复制控制描述符

该ARIB描述符（如第13章中所讨论）表示复制信息，包括自由复制、一次复制和禁止复制。

对于通过在数字复制控制描述符中的digital_recording_control_data进行复制限制的内容，或通过在内容可用性描述符中的encryption_mode进行复制保护的内容，接收器被禁止传输这样的内容

到任何可能使内容在因特网上传播的输出端。

23. 下载内容描述符

该ARIB描述符传递诸如大小、类型和ID的下载特征信息。

24. 紧急信息描述符

该ARIB描述符（如第13章中所述）用于广播紧急信息。

25. 事件组描述符

当多个事件之间存在相关性时，该ARIB描述符表示这些事件在一个组内。

26. 扩展广播商描述符

该ARIB描述符定义了扩展广播商标识信息，并且定义了这些广播商与其他扩展广播商。

27. 扩展事件描述符

该ARIB描述符提供了对事件的文本描述，它可以和*短事件描述符*一起使用。为了传递多于256字节的信息，可以采用多个描述符。DVB也使用该描述符。

28. 层次化传输描述符

该ARIB描述符（如第13章中所述）表示层次化数据流之间的关系。

29. 超链接描述符

该ARIB描述符描述了至其他事件、事件内容和信息事件的链接。

822

30. LDT链接描述符

该ARIB描述符描述了LDT中收集的信息的链接。

31. 链接描述符

该ARIB描述符（如第13章中所述）提供了至其他服务、传输流、节目指南、服务信息、软件更新等的链接。DVB也采用了该描述符。

32. 本地时间偏移描述符

该ARIB描述符可以存在于TOT中，用以描述与特定国家相关的相对于UTC的本地时间偏移的动态变换。它使得接收器可以自动在夏季和冬季时间之间进行调整。DVB也采用了该描述符。

33. logo传输描述符

该ARIB描述符描述了服务logo信息，如指向通过ARIB STD-B21传输的logo数据、logo标志符、logo版本和某个logo的字符串。字符串中的字母用8位表示。

34. Mosaic描述符

该ARIB描述符将数字视频分割成基本单元，并控制基本单元向逻辑单元的分配，提供逻辑单元的内容和对应信息之间的链接（例如业务、服务、事件等）。第13章对此进行了详细讨论。DVB也采用了该描述符。

35. 网络标识描述符

该ARIB描述符用于标识网络。

36. 网络名称描述符

该ARIB描述符以文本形式传输网络名，如“东京有线网络”。

37. 节点关系描述符

该ARIB描述符描述了两个节点之间的关系。

38. NVOD（准视频点播）参考描述符

该描述符，协同*时移服务描述符*和*时移事件描述符*，提供了一种有效地描述服务的方式，这些服务携带相同的事件序列，具有不同的起始时间偏移。

39. 家长分级控制描述符

该ARIB描述符提供了一种基于年龄的分类方式，并且提供了相关扩展以使得可以采用其他分类标准。第13章对此进行了详细讨论。DVB也采用了该描述符。

40. 局部接收描述符

该ARIB描述符描述通过地面传输路径的部分接收层次来传输的service_id。

41. 局部传输流描述符

SIT包含了用于控制、播放、复制局部传输流的所有信息。该ARIB描述符描述了这些信息。DVB也采用了该描述符。

42. 局部传输流时间描述符

该ARIB描述符描述了局部传输流的时间信息。

43. 参考描述符

该ARIB描述符表示来自节目和当地事件的节点参考。

44. 卫星传输系统描述符

该ARIB描述符用于传输卫星网络的物理参数，包括频率、轨道位置、东西标志、极性、调制模式和符率。DVB也采用了该描述符。

45. 系列描述符

823

该ARIB描述了一系列事件。

46. 服务描述符

该ARIB描述符以文本形式提供了服务和服务提供商的名称。DVB也采用了该描述符，只是ARIB和DVB的service_type_id不同。

47. 服务列表描述符

该ARIB描述符对每个传输流提供了一个服务和服务类型列表。DVB也采用了该描述符。

48. 短事件描述符

该ARIB描述符提供了事件名称以及对事件的短描述。DVB也采用了该描述符。

49. 短节点信息描述符

该ARIB描述符表示节点名称和对节点简单的解释。

50. SI参数描述符

该ARIB描述符表示SI参数。

51. SI Prime_TS描述符

该ARIB描述符表示SI基本TS的标识信息和它的传输参数。

52. STC参考描述符

该ARIB描述符表示当地事件的标识时间和STC之间的关系。

53. 流标识符描述符

该ARIB描述符使得流可以和EIT中的一个描述相关联。当一个服务中有多个流具有相同的类型时，该描述符非常有用。第13章对此进行了详细讨论。DVB也采用了该描述符。

54. 填充信息描述符

该ARIB描述符用于填充表格的内容，进行填充的原因可以是各种各样的。它也可以用于禁用已经失效的描述符。DVB也采用了该描述符。

55. 系统管理描述符

该ARIB描述符（如第13章中所述）表示所采用的广播和广播格式。

56. 目标区域描述符

该ARIB描述符（如第13章中所述）描述了一个事件的目标区域，或者描述构成一个事件的

流的部分。

57. 地面传输系统描述符

该ARIB描述符用于传输地面网络的物理参数，包括中心频率、带宽、卫星、层次结构、码率、防护间隔和传输模式。

58. 时移事件描述符

该ARIB描述符表示一个事件是另一个事件的时移后的副本。DVB也采用了该描述符。

59. 时移服务描述符

该ARIB描述符将一个服务链接到最多20个携带相同节目的服务，原节目和链接的节目在时间上有一个平移。典型的应用是准视频点播（NVOD）服务。DVB也采用了该描述符。

824

60. TS信息描述符

该ARIB描述符定义了分派给可用传输流的远程控制关键标志符，并且指示服务标志符和层次化传输中的传输层之间的关系。

61. 视频解码控制描述符

该ARIB描述符（如第13章中所述）控制基于MPEG的静态图像的解码。静态图像是通过低传输速率传输的，并且其目的是为了当视频编码方法发生变化时在视频结合点获得平滑的视频解码。

18.9 字幕

日本字幕数据（ARIB STD-B24第三部分）可以存在于视频PES、音频PES或独立的PES（优先选择）中。和视频无关的字幕称为“叠加”（ARIB STD-B5）。

字幕可以采用水平和垂直书写格式。支持的字符集包括Mosaic、中文、日本汉字、平假名、片假名、符号和字母数字。典型特征包括反转、闪烁、下划线、边缘线、阴影、粗体、斜体和粗斜体。同时也支持位图图形。

显示控制包括显示时序、消除时序、剪切、渐隐、擦除、滑动和滚动。它同时也支持灵活的视图、记录和播放选项。

18.10 数据广播

ARIB数据广播标准描述了在ARIB流中传输数据的可用的封装协议。基于MPEG-2 DSM-CC，它同样支持基于XML的多媒体编码策略。

目前，数据广播有5种不同的标准定义。每种标准定义都有附加的描述符以支持特定的标准。

18.10.1 数据轮转传输

这个规范传输常规的同步和异步数据，允许接收器在数据传输过程中获得数据。这种机制用于下载和多媒体服务。

18.10.2 数据管道

如果需要的话，该规范可以用于向接收器传输数据。数据直接在MPEG-2传输流包的载荷中携带。

18.10.3 事件信息传输

该规范用于向接收器中的应用程序发送同步和异步消息（即时发送或在特定的时间发送）。

它主要用于多媒体服务。

18.10.4 独立PES传输

825

该规范支持数据广播服务，这种数据广播服务采用以同步或异步方式进行的面向流的数据传输。数据在MPEG-2 PES包中携带。它同样用于字幕和叠加的字符。

18.10.5 交互信道协议

该规范提供了用于公共网络的传输协议，公共网络包括PSTN、ISDN和用于双向交互服务的移动网络。

18.11 应用框图

图18-2示出了一个典型的ISDB-S机顶盒结构图。

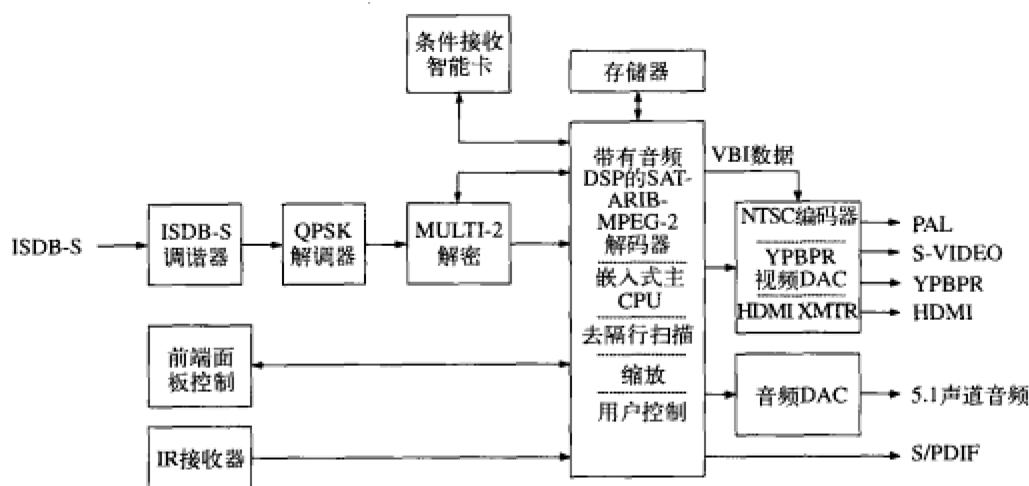


图18-2 ISDB接收器机顶盒框图

参考文献

1. STD-B10, *Service Information for Digital Broadcasting System*, version 3.8.
2. STD-B16, *Standard Digital Receiver Commonly Used for Digital Satellite Broadcasting Services Using Communication Satellite.*
3. STD-B20, *ISDB-S: Transmission System for Digital Satellite Broadcasting.*
4. STD-B21, *Receiver for Digital Broadcasting (Desirable Specifications)*, version 4.2.
5. STD-B23, *Application Execution Engine Platform for Digital Broadcasting.*
6. STD-B24, *Data Coding and Transmission Specification for Digital Broadcasting*, version 4.0.
7. STD-B25, *Conditional Access System Specifications for Digital Broadcasting.*
8. STD-B31, *ISDB-T: Transmission System for Digital Terrestrial Television Broadcasting*, version 1.5.
9. STD-B32, *Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting*, version 1.5.

826

10. STD-B40, *PES Packet Transport Mechanism for Ancillary Data.*

第19章 IPTV

随着数字视频和高速宽带网络的广泛应用，在宽带网络上传输实时音频和视频变得越来越流行。这种技术有几种叫法，包括IPTV（Internet Protocol TV）、流媒体、Video Over IP和IP视频。

这种技术并不是通过下载和存储大的视频和音频文件然后播放，而是数据在网络上以流的形式发送。流将视频和音频数据分割成适合在网络上传播的小的数据包。实时的视频和音频数据从视频服务器或实时视频编码器中流出，穿过网络，并在接收器（或客户端）上实时播放。因此，用户不必等到所有的数据都下载完就可以观看视频。

Telcos正在考虑将DSL和FTTH上实现的IPTV作为提供视频服务的一种方式以抗衡有线和卫星TV。他们现在已经可以提供VoIP（Voice over IP）、视频点播（VOD）、游戏、音乐、交互电视和当地的、全国的及付费的电视节目。

19.1 技术考虑

在网络上传输流视频并不是一件简单的事情。第一，即使是压缩后的视频数据，也需要相对较高的带宽。为了在一个1.5 Mbit/s的DSL链接上支持两个标清视频流，视频流的比特率应当限制在700kbit/s以内。因此，需要采用新的H.264和SMPTE VC-1视频压缩标准。H.264和VC-1所获得的低比特率可以使服务范围进一步扩大，因为DSL的比特率随着距离的增加而下降。

第二，为了避免播放过程产生中断，流视频数据需要实时传输。这就要求视频服务器和实时编码器能够连续地传输视频流并且可以避免网络拥塞。为了解决这一问题，已经有标准可以实现在网络上预留带宽资源。同时，还采用多播进一步减少对网络带宽的需求。

827

第三，流视频通常是突发式的。流视频客户端的接收缓冲的大小是有限的。如果不采取一定的措施平滑传输的比特率，接收缓冲就可能发生上溢或下溢。为了解决这一问题，需要采用另外的协议来管理时间问题。

19.2 多播

在网络上传送实时音频和视频的常用技术有以下三种。

单播，即服务器只将数据发送到一个接收器，如图19-1所示。端口号由接收器来选择。

广播，即数据从一个服务器向所有的接收器发送，如图19-2所示。

多播，即数据从一个服务器向一组接收器发送，如图19-3所示。服务器选择多点传送的IP地址和端口。这种方式是实况和准视频点播（NVoD）应用的一个特例。

近来，随着用户数量的增加以及实时传递大量数据（如视频、音频）的需要，多点传送逐渐得到了支持。在这种背景下，多点传送是节省网络带宽和服务能力的一种有效途径。

19.3 基于RTSP的解决方案

专用解决方案会使得每个服务提供商都有自己独特的流协议和文件格式，并且要求客户端能够支持多个协议或针对某一个特定的提供商。为了开发出一个开放的、标准的解决方案，IETF（Internet Engineering Task Force，因特网工程任务组）开发了几种协议以支持跨平台的互连和客户

与服务商之间的通信。

RTSP是一种控制协议，它可以初始化和引导来自视频服务商的流数据的传输，实现远程控制能力。RTSP并不传输多媒体数据，尽管RTSP连接可以用于传输RTP数据以方便防火墙和其他网络设备。

RTP是传输包括流音频和视频的实时数据的传输协议。RTP和RTSP通常一起使用，但也可单独使用。

RTCP是RTP的一部分，它可以用于辅助唇音同步和服务质量（Quality-of-Service，QoS）管理。

RSVP协议用于建立和管理所需的QoS级，确保有足够的网络资源（如带宽）可用。

19.3.1 RTSP

实时流协议（RTSP）在服务器（源）和客户端（接收器）之间建立和控制一个或多个时间同步的音频和视频数据流。服务器提供流的播放或录制服务，而客户端要求服务器提供连续的数据。

RTSP为音频和视频流提供"VCR风格"的控制功能，包括播放、暂停、快进、后退。它同时还提供以下功能。

828

从服务器获得节目信息。用户可以通过网络浏览器（HTTP）或其他技术请求可用节目的列表及其描述（节目描述）。如果一个节目是多点传送的，节目描述里也包含所用的多点传送地址和端口。如果一个节目只发送到一个客户端（单点传送），那么由用户端提供目的地址。

将服务接入会议。一个服务器可以被要求加入一个已有的会议，该服务器可以提供或记录数据。

向已有的数据中添加媒体。这对直播来说非常有用，当有附加的数据时，服务器可以即时通知客户。

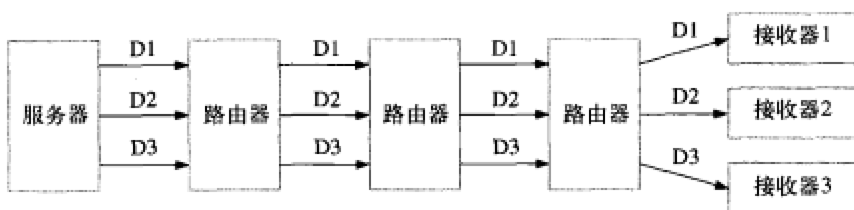


图19-1 单点传送举例。相同数据的三个副本以流D1、D2、D3的形式点对点地发送到接收端1、2、3

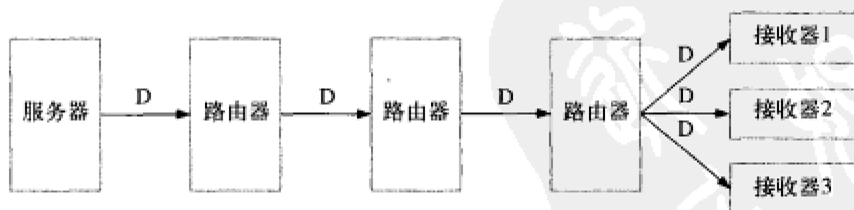


图19-2 广播举例。数据（D）的一个副本发送到所有的接收端

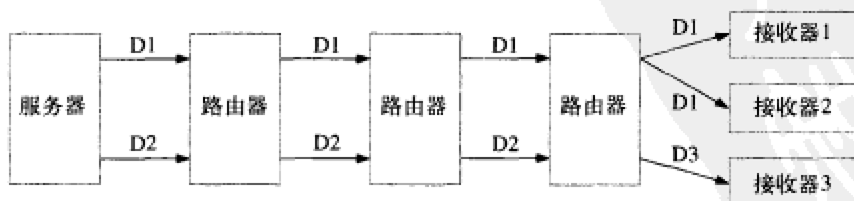


图19-3 多点传送举例。数据（D）的一个副本被多点传送到接收端1、2。注意，当接收端的个数增加时，这种方式可以节省局部和网际带宽

1. RTSP和HTTP的比较

RTSP和HTTP所提供的服务相同，只是RTSP是以音频和视频流的形式，而HTTP以文本和图形的形式。它们的语法和操作也都很相似，使得大多数的HTTP扩展也可以很容易地被RTSP采用。例如，RTSP的URL为rtsp://media.example.com:554/twister，表示节目twister，它可以由音频和视频流组成。RTSP的URL为rtsp://media.example.com:554/twister/audio，表示节目twister中的音频流，它可以通过向服务器media.example.com的端口554发出RTSP请求来控制。

由于用户界面通常采用Web页的形式实现，RTSP和HTTP的功能有一些重叠。因此，RTSP在Web和视频服务器之间支持不同的切换点。例如，节目描述可以通过HTTP或RTSP获得，从而使得RTSP服务器和客户端可以不支持HTTP。图19-4示出了采用Web服务器作为表现形式的实例，其中包括一个独立的视频服务器。

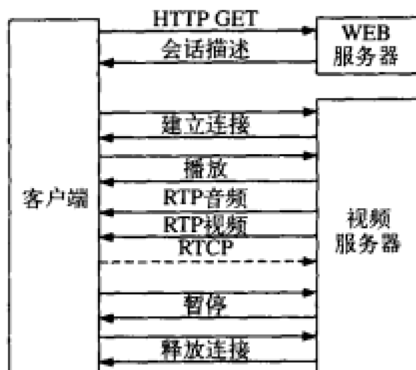


图19-4 客户端、Web服务器、视频服务器之间的通信

RTSP和HTTP的不同之处表现在两个主要领域。第一，不同于HTTP，RTSP兼容的视频服务器必须维持会话状态，以将RTSP请求和流关联起来。第二，从本质上说HTTP是一个不对称协议（客户端发出请求，服务器响应），但在RTSP协议中服务器和客户端都可以发出请求。例如，视频服务器可以发出请求以设置流的播放参数。

2. 流的特性

流的特性在*表现描述文件*中定义，其中包括编码格式、语言、RTSP URL、目的地址、端口和其他参数。客户端可以通过HTTP或其他方式获得表现描述文件。RTSP请求通常在独立于数据信道的信道上发送。

19.3.2 RTP

实时传输协议（RTP）是一种基于包的传输协议，它用来传输实时数据，如音频和视频流。RTP的设计初衷是针对多点传送，但也可以用于单点传送和视频点播。

829
830

在网络上发送数据包的延迟和误差是不可预测的，这使得实时传输视频变得更加复杂。为了解决这一问题，RTP包头包含时间戳、丢失保护、载荷标识、源标识和安全性等信息。这些信息用于在应用层实现数据包丢失恢复、拥塞控制等。

RTP通常运行于UDP的上层，以利用UDP的复用和求和校验功能。RTP是在两个主机之间提供基于连接的、稳定的数据流，而UDP则是在网络上提供一种非连接的（但是不稳定）数据报服务。RTP采用UDP作为目标传输协议的原因有两个。第一，RTP是针对多点传送设计的，基于连接的TCP不仅扩展性差，而且也不适合RTP。第二，对于实时数据，时间延迟比可靠性更重要，而TCP则采用重发机制实现较高的可靠性，这一点并不是RTP所需要的。例如，在网络拥塞的情况下，虽然一些包可能会丢失导致服务质量下降，但服务质量依然在可接收的范围内。如果协议坚持采用可靠的传输机制，重新传输的包可能会增加延迟、堵塞网络甚至饿死接收端。图19-5表示出了采用UDP/IP包封装的RTP包。



图19-5 采用UDP/IP包封装的RTP包

RTP本身并不提供保证时间延迟的机制。它要求下层提供支持以控制交换机和路由器中的资源。RSVP可以用于预留这种资源并提供请求的QoS。RTP也可以协同RTCP以获得关于数据传输质量的反馈信息以及会话参与者的相关信息。

RTP也可以协同RTCP以获得关于数据传输质量的反馈信息以及会话参与者的相关信息。

831 RTP被故意设计成一个不完整的协议框架。它向新的载荷格式和新的多媒体软件开放。通过增加新的档次和载荷定义，RTP可以很容易地适应新的数据格式和新的应用。

1. RTP会话

为了建立RTP会话，应用程序定义一对目的地址（一个网络地址加上两个针对TP和RTCP的端口）。在多媒体会话中，每种媒体通常采用其各自的RTP会话进行传播，对应的RTCP包报告该会话的传输质量。例如，音频和视频通常采用独立的RTP会话，允许接收器选择是否接收某个媒体。

2. 时间戳

时间戳对于实时应用来说非常重要。为了以正确的速率播放数据，接收端采用时间戳重建原始时间。时间戳也用于同步不同的流，如音频和视频数据。然而，RTP本身并不负责同步问题，同步是在应用层实现的。

此外，UDP并不按照时间顺序传输数据包。因此，序列号用于将接收的数据包按正确的顺序组织并实现包丢失保护。当一个视频帧被分割成几个RTP包时，一些视频格式允许所有的这些包具有相同的时间戳。因此，时间戳并不足以保证数据包可以被放回正确的顺序中。

3. 载荷标识

载荷标识符定义了内容的类型和压缩格式。它使得接收端知道如何解释、标识这些内容。支持的载荷类型有以下几种：

各种音频格式，包括CELP、线性PCM、ADPCM、G.711、G.712、G.722、Dolby® Digital、Dolby® Digital Plus、MP3等

MPEG-1音频和视频基本流

MPEG-1系统流

MPEG-2音频和视频基本流

MPEG-2程序和传输流

MPEG-4音频和视频流

MPEG-4 OD、BIFS、OCI和IPMP流

JPEG和M-JPEG视频流

DV（IEC 61834）、H.261和H.263流

SMPTE 421M（VC-1）视频流

MPEG-4.10（H.264）视频流

832 ASF

档次和载荷格式的定义可以通过增加额外的载荷类型来实现。在任何传输时间，RTP发送器只能发送一种类型的载荷，尽管载荷类型在传输过程中（例如为了适应网络拥塞）可以发生变化。

19.3.3　RTCP

实时控制协议（RTCP）是一种为了和RTP协同工作而设计的控制协议。在RTP会话中，客户端周期性地向服务器发送RTCP包，以传输数据传送质量的反馈信息和参加会话的成员信息。

RTCP包有以下5种类型。

接收端报告。接收端报告包含数据传输信息，包括最大接收包数、丢包数、抖动以及计算服务器和客户端之间的往返延迟的时间戳。

发送端报告。发送端报告包含接收端报告信息和与媒体间同步、累加包计数器和发送的字节数有关的相关信息。

源描述。包含描述源的信息。

Bye。表示参与过程结束。

针对特定应用的功能。当开发新的应用和新的特征时，用于实验研究。

通过这些控制信息包，RTCP可以提供以下功能。

QoS监测和拥塞控制。服务可以基于客户端的反馈调整传输。客户端可以决定拥塞是局部性、区域性还是全局性的。网络性能可以在多点传输过程中评估。

源标识。在RTP数据包中，源通过一个随机产生的32位的标识符标识，这一点对用户来说很不方便。源描述包包含如用户名称、电话号码、e-mail地址等文本信息。

媒体间同步。用于媒体间的同步，如音频和视频间的唇音同步。

控制信息扩展。当会话的参与者增加时，必须采取一定的措施防止控制通信拥塞网络资源。RTP将控制通信限制在整个会话的5%以内。这是根据参与者数量调整RTCP的产生率实现的。

RTP和RTCP一起提供传输实时内容时必要的功能和控制机制。然而，RTP和RTCP本身并不负责更高层的任务，如装配和同步。高层任务是在应用级完成的。

833

19.3.4 RSVP

资源预留协议（RSVP）使得网络可以支持QoS。

RSVP用于为网络资源建立如带宽等的资源存储。当客户端为其数据流请求一个特定的QoS时，它用RSVP沿着网络路径向节点（路由器）传输它的请求。在每个节点，RSVP试图为该数据流预留资源。一旦资源存储建立，RSVP同时负责维持请求的服务级别。

资源预留请求并不需要一直传播到服务器。实际上，每个资源预留请求向上游传播直到它遇到针对相同数据流的另一个资源预留请求，然后同那个请求合并在一起。这种资源预留请求合并是RSVP的主要优点之一：可扩展性——大量客户加入多点传送不会明显增加数据传输。RSVP可以很容易地针对大的多点传送组进行扩展，平均协议开销随着参与者的增加而下降。

RSVP同时支持多点传送和单点传送，并且能够适应成员和路由的变化。为了利用当前的因特网路由算法的健壮性，RSVP采用底层路由协议以决定是否携带资源预留请求。由于路由会根据网络的变化改变路径，RSVP可以调制其资源预留请求以适应新的路径。

19.4 ISMA

因特网流媒体联盟（ISMA）是一个由苹果公司、思科公司、IBM、Kasenna、飞利浦和Sun联合成立的非盈利工业联盟。自从成立以来，它已经获得了广泛的工业支持。它的任务是推动和促进在因特网上的音频和视频流媒体开放架构的推广。

ISMA v1.0提供了在网络上传输最大比特率为1.5Mbit/s的音频和视频流的工具。它采用MPEG-4音频/视频压缩标准，并针对内容传输和控制采用IETF协议（RTP、RTSP和SDP）。

ISMA v1.0定义了两个层次化档次：档次0和档次1。档次1支持档次0所支持的所有工具，同时增加了一些其他工具。

19.4.1 档次0

档次0是针对在无线和窄带网络上向只有有限的音频和视频处理能力的设备传输音频和视频

流的。

视频采用MPEG-4.2 SP@L1（QCIF，176×144）。音频采用MPEG-4.3 HQ@L2。音频最多支持双声道，采样率最高为48kHz。最大比特率为128kbit/s。

19.4.2 档次1

834

档次1的目的是在宽带网络上传输音频和视频流，以向用户提供更好的视觉体验。

视频采用MPEG-4.2 ASP@L3（CIF，352×288）。音频采用MPEG-4.3 HQ@L2。音频最多支持双声道，采样率最高为48kHz。最大比特率为1.5Mbit/s。

19.5 在IP上实现广播

ARIB和DVB传输流也可以在宽带IP网络上传输。传输流的包被封装成RTP包，并通过IP多点传输发送到客户端。

对于DVB，这种方式称为DVB-IPI或数字视频广播－网络协议设施。不要将DVD-IPI和DVB-IP混淆，DVB-IP是在DVB上实现IP服务。

19.6 条件接收（DRM）

对于宽带IP网络，条件接收通常称为DRM（Digital Right Management，数字版权保护）。

早期的IPTV开发中采用的DRM解决方案在原理上类似于DVB的同密技术。MPEG解码芯片中包含一个嵌入式（通常是改进的）AES或3DES解密模块；ISO 7816智能卡向AES/3DES解扰电路提供解密的关键信息。

新一代DRM解决方案不采用智能卡。在MPEG解码器芯片中安全运行的软件代替了智能卡，这不仅降低了成本，而且安全性更高。软件DRM解决方案通常也包括在内容被接收和解密后对该内容的使用和再发行进行控制的能力。DRM的能力包括：

- 关闭/打开模拟和数字视频复制保护；
- 限制模拟视频输出的分辨率（受限图像）；
- 限制数字音频输出的采样率和大小；
- 禁止模拟和数字音频和视频输出。

参考文献

1. ETSI TR 102 033, *Digital Video Broadcasting (DVB): Architectural Framework for the Delivery of DVB-Services Over IP-based Networks*, April 2002.
2. ETSI TS 102 813, *Digital Video Broadcasting (DVB): IEEE 1394 Home Network Segment*, November 2002.
3. ETSI TS 102 814, *Digital Video Broadcasting (DVB): Ethernet Home Network Segment*, April 2003.
4. IETF RFC 2205, *Resource ReSerVation Protocol (RSVP)—Version 1 Functional Specification*, September 1997.
5. IETF RFC 2206, *RSVP Management Information Base Using SMIv2*, September 1997.
6. IETF RFC 2207, *RSVP Extensions for IPSEC Data Flows*, September 1997.
7. IETF RFC 2208, *Resource ReSerVation Protocol (RSVP)—Version 1 Applicability Statement: Some Guidelines on Deployment*, September 1997.
8. IETF RFC 2209, *Resource ReSerVation Protocol (RSVP)—Version 1 Message Processing Rules*, September 1997.

835

9. IETF RFC 2250, *RTP Payload Format for MPEG1/MPEG2 Video*, January 1998.
10. IETF RFC 2326, *Real Time Streaming Protocol (RTSP)*, April 1998.
11. IETF RFC 2327, *SDP: Session Description Protocol*, April 1998.
12. IETF RFC 2343, *RTP Payload Format for Bundled MPEG*, May 1998.

13. IETF RFC 2750, *RSVP Extensions for Policy Control*, January 2000.
14. IETF RFC 3016, *RTP Payload Format for MPEG-4 Audio/Visual Streams*, November 2000.
15. *RTP Payload Format for MPEG-4 Streams*, IETF Internet Draft, April 2001.
16. IETF RFC 3189, *RTP Plaload Format for DV (IEC 61834) Video*, January 2002.
17. IETF RFC 3266, *Support for IPv6 in Session Description Protocol (SDP)*, June 2002.
18. IETF RFC 3550, *RTP: A Transport Protocol for Real-Time Applications*, July 2003.
19. IETF RFC 3551, *RTP Profile for Audio and Video Conferences with Minimal Control*, July 2003.
20. IETF RFC 3640, *RTP Payload Format for Transport of MPEG-4 Elementary Streams*, November 2003.
21. IETF RFC 3984, *RTP Payload Format for H.264 Video*, February 2005.
22. IETF RFC 4184, *RTP Paylaod Format for AC-3 Audio*, October 2005.
23. IETF RFC 4425, *RTP Payload Format for Video Codec 1 (VC-1)*, February 2006.
24. IETF RFC 4587, *RTP Payload Format for H.261 Video Streams*, August 2006.
25. IETF RFC 4598, *Real-time Transport Protocol(RTP) Payload Format for Enhanced AC-3 (E-AC-3) Audio*, July 2006.

836



术 语 表

8-VSB　参见残留边带（Vestigial Sideband）。

A-VSB　为了更好地满足移动市场，并与DVB-H和DMB形成竞争，ATSC A-VSB将改善动态多路径跟踪，准许使用分层（层次式）调制，支持时分复用和帧分片（slicing）。为了增加单一频率网络（Single Frequency Network，SFN）中对地覆盖的面积，在A-VSB中，不同塔台之间广播信号的定时也将更容易。

AC-3　Dolby®Digital压缩音频的别名。

交流耦合（AC Coupled）　交流耦合就是让模拟视频数据通过一个电容器以消除所有的直流分量（DC offset），或消除视频信号线上的总电压电平（voltage level）。一种拾取视频信号的方式是利用交流耦合来消除直流分量，然后进行直流分量恢复（DC restoration）以增加已知的直流分量（就是我们所选择的直流分量）。交流耦合之所以重要的另外一个原因是，它能够消除大的（有害的）直流分量。

有效视频（Active Video）　包含图像信息的那部分视频信号。在显示器上，有效视频如果不是全部可见的话，也是大部分可见。

AFC　见自动频率控制（Automatic Frequency Control）。

AGC　见自动增益控制（Automatic Gain Control）。

alpha　见alpha通道和alpha混合。

alpha混合（alpha Mix）　alpha值是用于控制两个图像中样本点的混合（mixing或blending）。

新采样点＝（alpha）（采样点A的颜色）＋（1－alpha）（采样点B的颜色）

通常情况下，alpha为归一化的值，即取值范围为0~1。32位帧缓存实际表达的意思就是24位的颜色值和8位的alpha值。

AM　见幅度调制（Amplitude Modulation）。

AMOL　广播节目自动测量（Automated Measurement of Lineups）的简写。通常情况下，Nielsen盒子（Nielsen box）使用的这种480i VBI信号位于线20、22、284和284上。

幅度调制（Amplitude Modulation）　将数据编码到载波的一种方法，在这种调制方法中，载波的幅度与数据值大小成比例。

抗混叠滤波器（Anti-Alias Filter）　这种滤波器是一种带限低通滤波器，经过该滤波器后信号带宽的频率范围为小于二分之一采样率。

ARIB　无线电工商业协会（Association of Radio Industries and Businesses）的简写，ARIB为日本的一个标准化组织。ARIB提供了几个规范，这些规范形成了日本使用的ISDB数字电视系统的核心。

ARIB STD-B10　日本ISDB-S和ISDB-T数字电视服务信息规范。

ARIB STD-B20　日本ISDB-S（卫星）数字电视系统规范。

ARIB STD-B21　日本ISDB-S和ISDB-T数字电视接收系统规范。

ARIB STD-B24 日本ISDB-S和ISDB-T数字电视数据广播规范。

ARIB STD-B25 日本ISDB-S和ISDB-T数字电视访问控制规范。

ARIB STD-B31 日本ISDB-T（地面）数字电视系统规范。

ARIB STD-B32 日本ISDB-S和ISDB-T数字电视视频编码、音频编码和复用规范。

ARIB STD-B38 日本ISDB-S和ISDB-T数字电视家庭服务器规范。

伪影（Artifact） 在视频领域中，伪影为瑕疵、噪音、雪花、污点等。当图像中有伪影时，从视觉的观点来讲，就是图像出现了某种类型的问题。不要将该术语与没有对显示器进行正确调整的情况发生混淆。例如，如果色度控制设置错误，那么图像看起来就感觉很差，但这不是伪影。伪影是图像的某种实际损坏。

宽高比（Aspect Ratio） 图像的宽度与高度之比。显示器的宽高比通常为4：3或16：9两种。节目内容可能还具有其他的宽高比（如2.35：1），这使得在显示器上显示时，会出现Letterbox模式的效应。

异步（Asynchronous） 参考没有公共时钟或时序信号的电路。

ATSC 先进电视系统委员会（Advanced Television System Committee）。定义了美国的HDTV标准。其他的国家也有采用ATSC HDTV标准的。

ATSC A/52 为ATSC HDTV定义了Dolby® Digital和Dolby® Digital Plus音频压缩标准。

ATSC A/53 为美国定义了ATSC HDTV。

ATSC A/65 为ATSC HDTV定义了节目和系统信息协议（Program and System Information Protocol，PSIP）。

ATSC A/70 为ATSC HDTV的条件访问系统（conditional access system）定义了标准。

ATSC A/80 针对数字电视馈送与发布应用，定义了一个卫星传输ATSC数据的调制与编码标准。

ATSC A/81 为ATSC直接到户（Direct-to-Home，DTH）卫星广播系统描述了发射系统。

ATSC A/90 为ATSC定义了数据广播标准。

ATSC A/92 定义了IP协议多播会话传输和IP多播ATSC A/90数据广播标准的使用。

ATSC A/93 定义了同步数据元素和同步及异步事件的传输。

ATSC A/94 定义了数据应用参考模型（Data Application Reference Model，DARM），其中包括应用环境设备与ATSC A/90数据广播标准的绑定。

ATSC A/95 针对层次式命名空间、目录和文件的传输，定义了ATSC传输流文件系统（Transport Stream File System，TSFS）标准。构建于ATSC A/90数据服务分发框架之上。

ATSC A/96 定义了一组核心协议族，使得可以在ATSC电视环境中进行远程交互。

ATSC A/97 为ATSC定义了软件下载数据服务。

ATSC A/100 这种DTV应用软件环境（DASE）定义了一个软件层（中间件），它准许编程内容和应用可以在公共ATSC接收设备上运行。

ATSC A/101，ATSC A/102 为ATSC定义了先进公共应用平台（Advanced Common Application Platform，ACAP）。

音频调制（Audio Modulation） 对具有音频信息的音频副载波进行修改，使得它可以与视频信息混合与发送。

音频副载波（Audio Subcarrier） 调制有音频数据的特定频率。

自动频率控制（Automatic Frequency Control，AFC） 锁定并跟踪特定频率的技术。

自动增益控制（Automatic Gain Control，AGC） 不论输入幅度是多少，都具有固定输出

幅度的一种电路。

AVC　高级视频编解码（Advanced Video Codec）的简写，MPEG-4.10（H.264）视频编解码早期使用的一个名字。

AVCHD　高分辨率便携式摄像机规范，它使用MPEG-4.10（H.264）视频压缩标准。

AVS　音视频编码标准（Audio Video coding Standard）的英文缩写，该标准由中国建立。

后肩（Back Porch）　模拟视频波形的一部分，它位于线同步的后肩与有效视频开始之间。

带通滤波器（Bandpass Filter）　仅准许选定的频率范围通过的电路。

带宽（Bandwidth，BW）　一个电路响应和能够通过的频率范围。它也可以是信号的最高频率与最低频率之差。

频带分段正交频分复用（bandwidth segmented orthogonal frequency division multiplexing）BST-OFDM试图改善COFDM，采用的方法是在同样的复用中调制某个与其他OFDM载波不同的OFDM载波。因此可以对给定的传输通道进行分割，在不同的频率片断上进行不同的调制。

基带（Baseband）　当应用于音频和视频时，基带就是音视频信号，它们并没有调制到别的载波上（例如，被调制到通道3或4上的RF）。在DTV中，基带也可以指压缩的（未调制的）位流。

黑场信号（Black Burst）　黑场信号是一种复合视频信号，总体上，它是一种黑色的图像。在同步视频设备中使用黑场信号使视频输出保持对齐。黑场信号告诉视频设备场同步、行同步和色度同步时序。

黑色电平（Black Level）　该电平定义了特定视频系统中什么样的电平代表黑色。如果由于某种原因，视频信号低于该电平，那么就认为它比黑色还要黑（blacker-than-black）。可以说，模拟同步信号比黑色还要黑。

消隐（Blanking）　在CRT显示器上，扫描线从左边向右边移动，然后跳回左边，并再次完成整个线的扫描，然后沿着显示屏向下移动。当扫描线抵达右边缘时，将折回到左边缘，在此过程中，视频信号被隐藏，因此我们不能“看到”扫描光束从右边缘到左边缘的返回路径。为了对视频信号进行消隐，模拟信号电平被降低到消隐电平上，如果使用的黑色电平与消隐电平存在差别（pedestal），那么消隐电平就低于黑色电平。

消隐电平（Blanking Level）　这种视频信号电平是由系统在消隐发生的地方定义的。如果没有使用黑色电平与消隐电平的差别，或者虽然使用了黑色电平与消隐电平的差别，但消隐电平低于黑色电平，那么消隐电平可能就是黑色电平。

高光溢出（Blooming）　这是一种效应，当视频变得比白色还要白时，有时就会出现高光溢处，在这种情况下，原本清晰的一根细线条在屏幕上就会变得粗大和模糊。

缓冲段（Breezeway）　模拟视频波形处在线同步后肩与色同步开始之间的部分。

亮度（Brightness）　指的是由显示器发出的光有多亮，它是由视频信号的强度控制的。

BS.707　该ITU建议为PAL和SECAM视频标准规定了立体声规范（Zweiton和NICAM 728）。

BST-OFDM　频带分段正交频分复用（bandwidth segmented orthogonal frequency division multiplexing）的英文缩写。

BT.470　该ITU建议规定了世界范围内使用的各种NTSC、PAL和SECAM视频标准。SMPTE 170M也规定了在美国使用的（M）NTSC视频标准。BT.470已经取代了BT.624。

BT.601　该ITU建议规定了720×480（59.94Hz）、960×480（59.94Hz）、720×576（50Hz）

和960×576（50Hz）的4：2：2 YCbCr隔行视频标准。

BT.653 该ITU建议规定了世界范围内使用的各种图文电视标准。为480i和576i视频系统定义了系统A、B、C和D类型。

BT.656 针对专业视频设备之间传输4：3 BT.601 4：2：2 YCbCr数字视频，该ITU建议定义了并行接口（8位或10位，27MHz）和串行接口（270MHz）。

BT.709 该ITU建议规定了1920×1080 R′G′B′和4：2：2 YCbCr隔行和逐行16：9数字视频标准。使用的帧率可以为60Hz、59.94Hz、50Hz、30Hz、29.97Hz、25Hz、24Hz和23.976Hz。

BT.1119 该ITU建议为480i和576i视频信号定义了宽屏信令（Wide Screen Signaling，WSS）。对于576i视频系统，WSS可以出现在线23上。对于480i视频系统，WSS可以出现在线22和线285上。

BT.1124 该ITU建议针对NTSC和PAL定义了重影消除参考（Ghost Cancellation Reference，GCR）。

BT.1197 该ITU建议定义了PAL+标准，允许在标准的PAL传输系统上传输16：9的电视节目。

BT.1358 该ITU建议定义了720×480（59.94Hz）和720×576（50Hz）的4：2：2 YCbCr专业逐行视频标准。也可以参考SMPTE 293M。

BT.1618 该ITU建议规定了基于DV的数据率为25Mbit/s和50Mbit/s的音频、数据和压缩视频的数据格式。也可以参考SMPTE 314M。

BT.1620 该ITU建议规定了基于DV的数据率为100Mbit/s的音频、数据和压缩视频的数据格式。也可以参考SMPTE 370M。

BTSC 该EIA TVSB5标准定义了针对NTSC视频实现立体声的技术。一个FM副载波传输一个L+R信号，而一个AM副载波传输一个L−R信号。

色同步（Burst） 参见颜色同步（color burst）。

色同步门（Burst Gate） 这是一个信号，它告诉NTSC或PAL视频解码器色同步位于扫描线的什么位置。

B-Y 蓝色－减－亮度（blue-minus-luma）信号，也称为色差信号。当将该信号加在亮度（Y）信号上后，可以产生蓝色视频信号。

载波（Carrier） 一种特定频率的波形，其上调制有需要传输的数据。

CBR 固定比特率（constant bit-rate）的缩写。

CCIR 国际无线电咨询委员会（Comite Consultatif International des Radiocommunications 或 International Radio Consultive Committee）。CCIR已经不复存在，它已经被归入了其上级组织（ITU）。对于给定的CCIR xxx规范，参见BT.xxx。

CEA-608 美国隐藏字幕（closed captioning）和扩展数据服务（Extended Data Service, XDS）标准。其修订版B增加了复制代次管理系统（Copy Generation Management System）——模拟（CGMS-A）、内容分级（V-chip）、使用Text-2（T-2）服务的因特网统一资源定位（URL）、16位传输信号标识符和DTV PSIP数据传输。

CEA-708 DTV字幕标准。EIA CEB-8也对嵌入ATSC数据流内的CEA-608数据流的使用和处理提供了指导，还提供了增强CEA-708。

CEA-805 该标准规定了如何将CGMS和AMOL数据加载到各种模拟视频信号上去。

CEA-861 该标准规定了如何在HDMI中包含像宽高比和格式信息这样的数据。

CGMS-A 复制代次管理系统（Copy Generation Management System）—模拟。参见CEA-

608和CEA-805。

超级VCD（chaoji VideoCD） 超级VCD（Super VideoCD）的另外一种称呼。

色度（Chroma） 一种视频信号，它包含两个部分，这两部分组成了我们在显示器上所看到的画面：亮度部分和颜色部分。色度就是颜色部分。

色度带通（Chroma Bandpass） 在NTSC和PAL视频信号中，亮度（白色和黑色）和色度（颜色）信息是结合在一起的。如果需要对NTSC或PAL视频信号进行解码，那么就必须将亮度信息和色度信息分离开来。色度带通滤波器就是从视频信号中将亮度信号移除的一种滤波器，并且使色度信息相对完整。除了图像中亮度信号和色度信号发生重叠的地方以外，这种滤波器都工作得相当好，这意味着亮度信号和色度信号所填充的频率范围有重叠。在重叠区，滤波器不能区分二者，因而二种信号都得以通过。这可能导致滑稽的画面。当观众在看电视的时候，在画面中出现一个人，他正穿着箭尾形的夹克或穿着瘦小、条纹紧密的衬衫，并且长相英俊。在这样的区域中，观众可能会看到一个彩虹样的彩色效果在移动。到底发生了什么事情呢？这是因为视频解码器错误地将亮度当作了色度。由于亮度并不是色度，视频解码器并不能断定到底显示哪个颜色，因此其显示效果看起来就像彩虹模式。这个问题可以通过梳状滤波器（Comb filter）来克服。

色度同步（Chroma Burst） 参见颜色同步（Color Burst）。

色度解调器（Chroma Demodulator） 在NTSC或PAL视频信号通过Y/C分离器后，必须将颜色进行解码。这就是解调器应该做的工作。它从Y/C分离器中取出色度信号，并恢复两个微分信号（典型情况下为I和Q，或U和V）。现在就得到了亮度信号和两个色差信号，因此视频系统就可以确定在屏幕上显示什么颜色了。

色度键（Chroma Key） 这是将两个视频图像结合起来的一种方法。进行色度抠像（chroma keying）的一个例子是站在巨幅气象图前的晚间新闻主持人。实际上，这个主持人是站在蓝色或绿色的背景前面的，最后得到的图像是与计算机生成气象图的合成效果。下面就是它的工作原理：一台电视摄相机（TV camera）对准主持人和气象图，同时气象图也和该主持人一起传入电视摄像机（box），摄相机需要进行决策。在看到蓝色或绿色背景的任何地方，就显示气象图。否则就显示主持人。因此，任何时候只要主持人在附近移动，那么摄像机就需要确定主持人所在的位置，并显示正确的画面。

色度陷阱（Chroma Trap） 在NTSC和PAL视频信号中，亮度（白色和黑色）和色度（颜色）信息是结合在一起的。如果需要对NTSC或PAL视频信号进行解码，那么就必须将亮度信息和色度信息分离开来。色度陷阱就是将色度信息与亮度信息分离开的一种方法，相对来讲，它使得亮度信息原封不动地得以保留。它是如何工作的呢？将NTSC和PAL信号输入陷波滤波器。对于所有实际的应用，陷波滤波器只准许特定频率的信号得以通过，而其他频率的信号则无法通过。设计具有一定频率响应的陷波滤波器是为了移除色度信号，这种滤波器的输出就仅包含亮度信息了。由于该滤波器去掉了色度信息，因此这种滤波器被称为陷波滤波器。在所有这些操作中，令人遗憾的部分是，这种滤波器不仅移除了色度信息，如果在陷阱所在的频率范围内同时存在亮度信号，那么，它们也会一起被滤掉。滤波器只知道它所要滤去信号的频率范围，对于不同的图像，亮度信息可能与色度信息发生重叠。滤波器不能区分亮度和色度之间的差别，因此，当二者在同一频率范围内时，陷阱滤波器就会将它们都滤掉。更重要的是什么呢？喔，将失去亮度信息，这意味着图像质量在某种程度上变差了。使用梳状滤波器（comb filter）作为Y/C分离器要优于色度陷阱或色度带通滤波器。

色度（Chrominance） 在视频领域中，术语色度（Chroma）和色度（Chrominance）通常（错误地）互换使用。参见色度（Chroma）的定义。

CIF 公共接口格式（Common Interface Format）或公共图像格式（Common Image Format）。其支持的有效分辨率为352×240或352×288。CIF格式的变种包括2CIF（支持的分辨率为704×240或704×288）、DCIF（支持的分辨率为528×320或528×384）、4CIF（支持的分辨率为704×480或704×576）和HD-CIF（支持的分辨率为1920×1080）。

箝位（Clamp） 这基本上是DC恢复电路的另外一个名字。它也可以指用于DC恢复电路中的开关电路。当表示DC恢复时，它通常用作clamping，当表示开关时，通常是clamp。

裁剪逻辑（Clipping Logic） 用于防止无效变换的电路。某些颜色可以出现在一个颜色空间中，但却不能出现在另外一个颜色空间中。就在从一个颜色空间转换到另外一个颜色空间后，颜色空间变换器可能需要检查出无效的颜色。如果出现任何无效的颜色，那么裁剪逻辑就需要对其进行限制，或称为剪切掉这一部分信息，直到出现合法的颜色值为止。由于这种电路剪切掉了某些信息，并且是用逻辑来实现的，这就不难看出，名字“裁剪逻辑”是如何得来的啦。

隐藏字幕（Closed Captioning） 字幕是一种服务，它对与视频信号一起发送的文本信息进行解码，并将其显示在显示器的屏幕上。有两种主要的字幕规范CEA-608和CEA-708。

隐藏字幕（Closed Subtitles） 参见字幕（Subtitles）。

编码正交频分复用（Coded Orthogonal Frequency Division Multiplexing） 编码正交频分复用（COFDM）。它以不同于8-VSB或其他单载波方法来传输数字数据。频分复用（Frequency Division Multiplexing）的意思是要发送的数据是分布在多个载波（对于DVB-T为1705或6817个载波）上，这是与调制到单一载波上的情况相对应的。这样，在每个COFDM载波上的数据率要远低于单个载波所要求的数据率。COFDM载波是正交的，或者说互相垂直的，并且使用了前向错误校正（编码的）。

COFDM是一种复用技术，而不是一种调制技术。任何一种通常的调制技术都可以用于调制COFDM载波，如QPSK、16-QAM或64QAM。

COFDM 参见编码正交频分复用（Coded Orthogonal Frequency Division Multiplexing）。

彩条（Color Bar） 这是一种用于检查视频系统是否得到正确校正的测试模式。如果颜色具有正确的亮度、色度和饱和度，那么该视频系统就得到了正确校正。

色同步（Color Burst） 一种特定频率和幅度的模拟波形，它位于线同步信号的后肩和有效视频的起始之间。色同步告诉NTSC或PAL视频解码器，如何对包含在该线有效视频中的颜色信息进行解码。通过观察色同步，解码器就可以确定什么是蓝色、橙色或洋红色。本质上讲，解码器就是要确定什么是正确的颜色。

颜色解码器（Color Decoder） 参见色度解调器（Chroma Demodulator）。

颜色解调器（Color Demodulator） 参见色度解调器（Chroma Demodulator）。

颜色微分（Color Difference） 用于颜色视频中的所有颜色空间都需要3个分量。它们可能是R′G′B′、YIQ、YCbCr、YUV或Y（R－Y）（B－Y）。在Y（R－Y）（B－Y）颜色空间中，基于显而易见的理由，R－Y和B－Y分量经常被称为色差信号。它们是通过将亮度（Y）减去红色和蓝色分量得到。I和Q、以及U和V也是色差信号，因为它们是R－Y和B－Y的缩放版本。尽管在SDTV和HDTV之间，YIQ、YUV和Y（R－Y）（B－Y）有细微的差异，但其中的每个Y值基本上是一样的。

颜色镶边（Color Edging） 出现在物体边缘的额外颜色，但这些颜色与物体所在的区域并没有关系。

颜色编码器（Color Encoder） 颜色编码器的工作与颜色解码器刚好相反。它以两个微分信号作为输入，如I和Q，或U和V，并对它们进行联合处理以得到色度信号。

颜色键（Color Key） 本质上与色度键（Chroma Key）是一样的。

消色器（Color Killer） 消色器是一种电路。如果传入的模拟视频并不包含颜色信息，那么消色器就关闭颜色解码。这个过程是如何工作的呢？消色器查找色同步信息，如果找不到，它就关闭颜色解码器。例如，让我们看一看，彩色电视机接收采用黑白方式记录的内容，由于黑白信号并不包含色同步，颜色解码器就被关闭了。为什么要使用消色器呢？起初，颜色解码器还只能生成很少位数的颜色，如果接收到的是黑白传输信号，颜色解码器的微小错误就可能导致黑白节目中图像只有微弱的色花（color spot）。

颜色调制器（Color Modulator） 参见颜色编码器（Color Encoder）。

颜色纯度（Color Purity） 该术语用于描述实际颜色值与理论值的接近程度。例如，在YUV颜色空间中，颜色纯度规定为饱和度和±q的百分比，其中q为以度为单位的角度，两个量都是以兴趣颜色（color of interest）为参考的。这两个数越小，实际值就与真正假定的值越接近。对于演播室级（studio-grade）的设备而言，饱和度为±2%，色度为±2度。在矢量显示器上，如果颜色在这个范围内，就达到了演播室级水平。

颜色空间（Color Space） 颜色空间是颜色的一种数学表示。不管使用什么样的颜色空间（R′G′B′、YIQ、YUV等），橙色依然是橙色。变化的只是怎样去表示橙色。例如，R′G′B′颜色空间是基于笛卡儿坐标系统的，而HIS颜色空间则是基于极坐标系统的。

颜色副载波（Color Subcarrier） 颜色副载波是一种用于对颜色NTSC/PAL编码器或颜色解码器进行控制的信号。对于（M）NTSC而言，颜色副载波的频率大约为3.58MHz，而对于（B、D、G、H、I）PAL而言，它大约为4.43MHz。在颜色编码器中，一部分颜色副载波用于创建色同步，而在颜色解码器中，色同步用于重建颜色副载波。

色温（Color Temperature） 色温以开氏温度（Kelvin）的度进行度量。如果电视机的色温为8000开氏度，这意味着，白色与加热到这个温度的一片纯炭具有相同深浅的颜色（shade）。低色温向红色偏移，高色温向蓝色偏移。

对于视频来说，色温的标准为开氏温度6500度。这样，专业电视机的显示器使用6500度的色温。然而，大多数消费类电视机具有的色温为8000度或更高。这导致向蓝色投影。通过调整电视机的色温，可以产生大多数的准确颜色，但这是以损害图像的亮度为代价的。

梳状滤波器（Comb Filter） 这是另外一种进行NTSC和PAL的Y/C分离的方法。梳状滤波器用于代替色度带通滤波器或色度陷阱。梳状滤波器提供了更好的视频质量，因为它能够更好地将亮度信号与色度信号分离。它减少了拉链纹伪影（creepy-crawlies 或 zipper artifact）的数量。之所以称之为梳状滤波器，是因为其频率响应看起来像一把梳子。需要记住的一个重要事情是，对于Y/C分离来说，梳状滤波器这种方法要优于色度带通滤波器和色度陷阱。

公共图像格式（Common Image Format） 参见CIF。

公共接口格式（Common Interface Format） 参见CIF。

分量视频（Component Video） 使用三个独立颜色分量的视频，如YCbCr（数字的）、YCbCr（模拟的）或R′G′B′（数字的或模拟的）。

复合视频（Composite Video） 单一的模拟视频信号，其中包含亮度、颜色和定时信息。如果一个视频系统可以正确地接收视频，那么它必须拥有处于适当位置的多种信息。它必须具有用于在显示器上显示的图像，并且必须能够以正确的颜色进行显示。这类信息称为有效视频。视频系统还需要一些告诉它在什么地方显示每个像素的信息，这类信息称为同步。显示器需要知道在什么时候关闭电子束，使得观看者不能看到光束在CRT显示器上回折的过程。这类视频信息称为消隐。现在，每类信息都可以在三个独立的连接通道上并行传输。它仍然可以称为视频，并且

在显示器上看起来也很好。但这是一种浪费，因为，所有这三类信息都可以结合在一起，使得仅需要一个连接通道来发送。复合视频就是这样的视频流，它将显示一幅图像所需的所有类型的信息都结合成一个信号，这样，就仅需要一个连接通道。NTSC和PAL就是复合视频的例子。二者都是由有效视频、线同步、行消隐、场同步、场消隐和色同步组成的。尽管每个红、绿和蓝信号都可以包含同步和消隐信息，但R′G′B′并不是复合视频的例子。这是因为为了以正确的颜色在屏幕上显示画面，需要所有的三个信号。

压缩比（Compression Ratio）　压缩比是用于表示对一幅图像进行压缩时，其中有多少信息被压缩的一个数值。例如，假定我们开始有一幅1MB大小的图像，然后将其压缩为128KB。那么压缩率为：

$$1\,048\,576/131\,072 = 8$$

这表示压缩比为8 : 1；压缩后的图像只需要原来图像存储空间的1/8了。对于给定的压缩技术（例如MPEG），压缩比越高，那么图像看起来就越差。这与哪种压缩方法更好是没有关系的，例如，JPEG与MPEG相对比。使用MPEG以100 : 1的压缩比压缩过的视频流，可能看起来要好于使用JEPG对同样的视频流以100 : 1的压缩比进行压缩的结果。

条件接收（Conditional Access）　这是一种技术，通过这种技术，服务提供商就可以使购买该服务的用户对视频内容进行解码和观看。这种技术由密码解密（使用来自于随内容周期发送的、不断变化的编码密码）和解扰组成。解密技术可能是具有版权的（如Canal＋、DigiCiper、Irdeto Access、Nagravision、NDS、Viaccess等），或是标准化的（如DVB公共加扰算法和OpenCableTM）。条件接收可以认为是数字版权管理的一个简单形式。

两个公共DVB条件接收（CA）技术为SimulCrypt和MultiCrypt。对于SimulCrypt而言，一个单一的传输流可以包含多个CA系统。这使得具有不同CA系统的接收设备可以接收和正确地解码同样的音视频流。对于MultiCrypt而言，接收设备准许用户手动在CA系统之间进行切换。这样，当提供给观看者的CA系统并没有安装在接收设备中时，观看者只需要简单地切换一下CA卡。

固定比特率（Constant Bit Rate，CBR）　表示位流（压缩的或未压缩的）在传输过程中，每秒钟传输的位数是一样的。

轮廓（Contouring）　这是一种图像伪影，它是由于没有足够的位数来表示图像而引起的。之所以称之为“轮廓”是因为，图像的边缘在垂直方向上建立了亮度带。

对比度（Contrast）　视频术语，指的是在视频波形中，最白的白色与最黑的黑色相差有多远。如果峰值白色与峰值黑色相差很远，那么就说该图像具有高的对比度。具有高对比度，那么图像看起来就颜色分明（stark），并且相对来讲，非常像黑白的铺有瓷砖的地面。如果二者之间非常靠近，那么就说图像具有差的或低的对比度。图像具有低对比度，那么可以说该图像被洗刷过（washed out）：无法区分黑色和白色的差异，图像看起来就显得灰暗。

拉链纹伪影（Creepy-Crawlies）　是的，这才是一个真正的视频术语！拉链伪影指的是一种特定的图像伪影，它是由NTSC系统造成的。当在播放晚间新闻时，包含一副图像的矩形区域出现在新闻节目主持人的肩膀上方，或当某些计算机生成的文字出现在播放的视频剪辑上面时，靠近电视屏幕，并仔细观察，将发现有些图像具有向上（或向下）翻转的锯齿。这就是拉链伪影。有的人称之为拉链伪影（Zipper）就是因为其看起来像拉链。

串色（Cross Color）　当NTSC或PAL视频解码器错误地将高频亮度（luma或brightness）信息解释为色度（chroma或color）信息时，就会出现串色这种现象。这会导致颜色被显示在错误的位置。

串亮度（Cross Luma） 当NTSC或PAL视频解码器不正确地将色度（chroma或color）信息解释为高频亮度（luma或brightness）信息时，就会出现串色这种现象。

交叉调制（Cross Modulation） 当一个信号错误地调制了另一个信号时所要满足的一个条件。

串扰（Crosstalk） 在一个信号中检测到的来自另一个信号的干扰。

CVBS 复合视频基带信号（Composite Video Baseband Signal）的缩写，或复合视频、消隐和同步（Coposite Video, Blanking and Synchronization）的缩写。

DAVIC 数字音视频委员会（Digital Audio Visual Council）的简写。其目标是为广播、交互数字音视频信息和多媒体通信建立端到端互操作的企业标准。目前形成的这些规范有ISO/IEC 16500（标准化的部分）和ITR 16501（非标准化的部分）。

dB 为分贝（decibel）的缩写，表示相对能量（如电压或电流）的标准单位。

dBm 用增加或减少值的对数关系表示的通信中的能量测度，0dBm=1mW。在50Ω的系统中，0dBm=0.223V。

dBw 以1瓦为参考的分贝值。

直流分量恢复（DC Restoration） 在对模拟视频信号进行交流耦合后，必须对模拟视频信号进行直流分量恢复并进行数字化。由于视频波形进行了交流耦合，我们不再知道视频波形的绝对位置了。例如，同步提示的底部在–5V处还是在1V处呢？事实上，不仅我们不知道它位于什么地方，即便知道了，它还是随时间变化的，因为有效视频的平均电压值是随时间变化的。由于ADC要正常工作，要求具有已知的输入电平和范围，视频信号需要参考已知的直流分量电平。本质上，直流分量恢复就是在交流耦合信号上加上一个已知的直流分量电平。在对视频进行解码时，用于直流分量恢复的直流分量电平通常就是这样一个值，直流分量电平要使得同步提示在数字化后生成的数字为0。

DCAS 可下载条件接收系统（Downloadable Conditional Access System）的简写。

DCT 它是离散余弦变换（Discrete Cosine Transform）的缩写，很多视频压缩算法中都使用了离散余弦变换。

分贝（Decibel） 为十分之一贝尔（Bel），用于根据增益或损失对两个能量（如电压或电流）之比进行定义。它定义为10倍能量比，或20倍电压或电流比以10为底的对数（lg）值。

抽取（Decimation） 当模拟信号被数字化以后，假定产生了100个样本，但每两个样本只有一个可以保存或使用，这样得到信号以因子2：1抽取。得到的图像只有原来图像大小的1/4，因为3/4的数据被丢弃了。如果仅使用了图像大小的五分之一的数据，那么图像的抽取因子为5：1，抽取后的图像只有原来图像大小的1/25。因此，抽取是一种又快又容易的图像缩放方法。

抽取可以用多种方式来实现。一种方式就是上面描述的方法，其中，数据在表面上被直接丢弃。尽管这种技术容易实现，并且廉价，但它会引入走样伪影（aliasing artifacts）。另一种实现方式是使用抽取滤波器（decimation filter），它减少了走样伪影，但实现代价比较大。

抽取滤波器（Decimation Filter） 抽取滤波器是一种低通滤波器，是设计用于提供抽取功能的，但却不具有简单丢弃数据而产生的走样伪影。

去－加重（De-Emphasis） 也可以参考后－加重（Post-Emphasis）和后－均衡（Post Equalization）。去－加重具有的频率响应特征是要对预－加重（pre-emphasis）引入的特征进行补偿。

去－加重网络（De-Emphasis Network） 去－加重网络是一种电路，用于将频率响应恢复到其初始形式。

深色（Deep Color） 在每个颜色分量用10位或12位数据来表示数字YCbCr或R′G′B′视频数据的情况下，就称这样的颜色为深色。

解调（Demodulation） 从调制载波中恢复原始信号的过程。

解调器（Demodulator） 在NTSC和PAL视频中，解调是用于恢复色差信号的技术。参见色度解调器和颜色解码器的定义，在NTSC和PAL视频应用中，它们是解调器的另外两个名称。解调也经常置于DTV调频器之后，用于将传输的DTV信号转换为压缩位流。

微分增益（Differential Gain） 微分增益是，当亮度电平发生变化（并不希望发生）后，颜色饱和度变化的幅度。对于视频系统，微分增益越好（即指定的值越小），系统在确定正确颜色方面的能力就越强。

微分相位（Differential Phase） 微分相位是亮度发生变化（并不希望这样）后，色度变化的多少。对于视频系统，微分相位越好（即指定的值越小），系统在确定正确颜色方面的能力就越强。

数字分量视频（Digital Component Video） 使用三个独立颜色分量的数字视频，如YCbCr或R′G′B′。

数字复合视频（Digital Composite Video） 本质上讲是NTSC或PaL视频信号波形的数字化，并赋予同步、消隐和白色电平以特定的数字值。

数字版权管理（Digital Rights Management，DRM） 数字版权管理（DRM）是很多能力的一个总称，这些能力准许视频内容生产者或发布者来决定在什么条件下他们的产品可以获取、存储、观看、复制、租赁等。

数字传输内容保护（Digital Transmission Content Protection，DTCP） 一种加密方法（也是我们所知道的5C或DTCP），它是由索尼、日立、Intel、松下和东芝公司为IP、USB和IEEE1394接口制定的。

数字通用磁盘（Digital Versatile Disc，DVD） 参见DVD视频（DVD-Video）和DVD音频（DVD-Audio）。

离散余弦变换（Discrete Cosine Transform，DCT） DTC只是表示图像的另外一种方式。它不是从时域来看待图像（顺便讲一下，这就是我们通常看待图像的方式），而是在频域中来看待图像。它类似于颜色空间，在颜色空间中，颜色还是那个颜色，但是表示方式却不一样了。在此，也是一样，图像还是那个图像，只是表示图像的方式不同而已。

为什么JEPG、MPEG、H.261和H.263的部分压缩策略都是建立在DCT的基础上呢？这是因为，以这种方式来表示图像更加高效。同样，用YCbCr颜色空间来表示图像要比用RGB来表示图像更加高效，DCT在图像表示方面更加高效得多。

离散时间震荡器（Discrete Time Oscillator，DTO） 离散时间震荡器是压控振荡器的数字版本。

dNTSC™ 由Dotcast公司开发的用于在NTSC视频信号内进行数据广播的一种技术。

Dolby® Digital Dolby® Digital Plus由Dolby®公司建立的音频压缩技术。二者都是用于DVD、HD DVD、蓝光（Blue-Ray）和DTV的多通道环绕声格式。

点距（Dot Pitch） 测得的屏幕像素之间的距离，以毫米（mm）表示。这个数值越小，水平分辨率就越好。

双缓冲（Double Buffering） 正如其名，正在使用两个缓冲区。对于视频来讲，这意味着有两个帧缓冲。当缓冲区1正在进行读操作时，缓冲区2则正在进行写入操作。当二者的操作完成以后，缓冲区2则进行读出，而缓冲区1则进行写入。

下变频器（Downconverter） 这是一种用于将高频信号变为低频信号的电路。

下行链路（Downlink） 卫星用于发送数据到地面站的频率。

DRM 参见数字版权管理（Digital Rights Management）。

DMB 数字多媒体广播（Digital Multimedia Broadcasting）的缩写，由韩国建立的。它是通过卫星（DMB-S）或地面（DMB-T）进行广播，并将MPEG-4.10（H.264）用于视频编码，而将MPEG-4.3 BSAC或HE-AACv2用于音频编码。将音视频封装在MPEG-2传输流中。DMB是一个ETSI标准（TS 102 427和TS 102 428）。

下降域加扰（Drop Field Scrambling） 除了没有对行消隐期进行抑制外，这种方法与用于对模拟TV信道加扰的同步抑制技术是一样的。同步脉冲抑制仅出现在场消隐期间。在场消隐期内，解扰脉冲仍然会出现（这就愚弄了非授权的解扰设备）。如果解扰设备仅被解扰脉冲所触发，那么，该设备就不知道加扰器正在使用下降域加扰技术。这就是所谓的重插入（reinsertion），它使得有效视频信号得到了压缩。一个未授权的解扰设备只能建立冲刷（washed out）图像，并损失了下降域加扰期间的中性同步信号（neutral sync）。

DCTP 数字传输内容保护（Digital Transmission Content Protection）的缩写。

DTS® DTS®代表数字影院系统（Digital Theater System）。它是多通道环绕声格式，类似于Dolby® Digital。对于使用了DTS®音频的DVD，DVD视频规范仍然需要PCM或仍然需要提供Dolby® Digital音频。在此情况下，仅可以提供两个Dolby® Digital音频通道（由于带宽的限制）。

DTV 数字电视（Digital Television）的缩写，包括SDTV、EDTV和HDTV。

DV 数字视频（Digital Video）的缩写，这是一个用于在磁带上记录数据的数字便携式摄像机（camrecorder）标准。它是由BT.1618、BT.1620、SMPTE 314M和370M规范来定义的。

DVB 数字视频广播（Digital Video Broadcasting）的缩写，它是用于传输数字音视频的一种方法。有多个变种：DVB-T，这是用于地面广播的；DVB-S，这是用于卫星广播的；DVB-C，这是用于有线广播的；DVB-H和DVB-SH，这是用于手持设备的；DVB-IP是用于IPTV的。

DVD音频（DVD-Audio） DVD包含线性PCM音频数据，其采样率可以是44.1kHz、48.0kHz、88.2kHz、96.0kHz、176.4kHz或192kHz的任意组合，每个样本可以是16、24或32位数据，可以有1~6个通道，最大比特率可达9.6Mbit/s。只有两个通道准许176.4kHz或192kHz的采样率。

无损压缩编码（Meridian Lossless Packing，MLP）为一种无损压缩方法，其压缩比大约为2：1。MLP压缩是可选的，但在所有的DVD音频播放器中，解码能力是强制的。

对于DVD音频磁盘的任何视频部分来讲，都要求具有Dolby® Digital压缩音频。

DVD视频（DVD-Video） DVD包含大约2小时的数字音频、视频和数据。视频是使用MPEG-2 MP@ML进行压缩和存储的。使用的是可变比特率，平均比特率大约为4Mbit/s（仅适用于视频），峰值比特率为10Mbit/s（仅适用于视频）。音频要么是采用线性PCM或Dolby® Digital压缩音频。DTS®压缩音频也可以作为一个选项使用。

线性PCM音频的采样率可以是48kHz或96kHz，每个样本可以是16、24或32位数据，1~8个通道。最大比特率为6.144Mbit/s，这限制了每种情况下的采样率和位深度。

对于Dolby® Digital，比特率可以为64~448kbit/s，5.1通道的正常比特率为384kbit/s，立体声的正常比特率为192kbit/s。通道组合可以为（前/环绕）1/0、1＋1/0（双单声道）、2/0、3/0、3/1、2/2和3/2。LFE通道（0.1）是可选的，它具有所有8种组合。

对于DTS®音频，位率为64~1536kbit/s。通道组合（前/环绕）：1/0、2/0、3/0、2/1、2/2、3/2。LFE通道（0.1）为可选的，它具有所有6种组合。

DVI 数字视频接口（Digital Video Interface）的缩写。这是一种与显示器的视频接口，设

计这种接口主要是为了替换模拟YCbCr或R'G'B'接口。对于模拟显示器，D/A转换过程在显示器中进行。CEA-861标准规定了如何包含像宽高比和数据格式之类的信息。VESA EEDID和DI-EXT标准建立了数据结构和跨DVI进行数据通信的机制的相关文档。

DVI-D是一个纯数字的连接头；DVI-I连接头可以同时处理模拟和数字信号。DVI-A可用于插（公）连接头并实现与纯模拟的DVI-I连接头针脚的连接。DVI-A仅用于适配器电缆，在适配器电缆中，需要将数字信号转换为传统的模拟VGA信号，或由传统模拟VGA信号转换为数字信号。

DVITC 数字场消隐期时间码（Digital Vertical Interval Timecode）的缩写。

动态范围（Dynamic Range） 电路可以接收或生成的最弱到最强的信号。

E-AC-3 Dolby® Digital Plus压缩音频的另一个名称。

E-VSB 第二个更可靠的ATSC通道。

EDTV 参见增强分辨率电视（Enhanced Definition Television）。

EIA 电子企业联盟（Electronics Industries Alliance）的缩写。

EIA-516 美国使用的图文标准，也称为NABTS。

EIA-744 定义了V-chip（V型芯片）操作。该标准在CEA-608字幕标准中增加了内容分级过滤（content advisory filtering）能力。现在包含到了最新的CEA-608标准中，而EIA-744已经被撤销了。

EIA-761 规定了如何将QAM转换为8-VSB，支持OSD（在屏幕显示器上）。

EIA-762 规定了如何将QAM转换为8-VSB，不支持OSD（在屏幕显示器上）。

EIA-766 美国HDTV内容分级标准。

EIA-770 该规范由三部分组成。EIA-770.1和EIA-770.2为480i系统定义了模拟YCbCr视频接口。EIA-770.3为HDTV系统定义了模拟YCbCr视频接口。CEA-805定义了如何在这些YCbCr视频接口上传输VBI数据。

EIA-775 针对连接到使用IEEE 1394标准的DTV，EIA-775定义了一个基带数字接口规范，并提供了类似于模拟系统的功能级别。EIA-775的设计是为了实现DTV和各种类型的消费类数字音视频源之间的互操作，这些音视频源包括机顶盒和DVR或VCR。

EIA-775.1增加了一种机制，使得准许MPEG服务源可以在DTV上利用MPEG进行解码和显示的能力。

EIA-775.2增加了一些关于如何在DTV或其他的源设备上使用数字存储设备的信息，存储设备是DVHS或硬盘数字记录设备等类似的这类设备，而源设备则为用于记录的有线电视机顶盒，或时漂（time-shift）数字电视信号等。该标准支持如下这类设备的使用，这些设备是通过如下一些定义来进行使用的：服务选择信息（Service Selection Information，SSI）、用于管理发生在记录或回放过程中的中断的方法，以及用于对局部传输流进行管理的规则。

EIA-849规定了EIA-775标准各种应用的档次（profile），包括与ATSC地面广播兼容的数字流、直播卫星（DBS）、OpenCable™和标清数字视频便携式摄像机。

EIA-J CPR-1204 该EIA-J建议为480i视频信号规定了另外一种宽屏信令（WSS）标准。WSS可以出现在线20和283上。

增强清晰度电视（Enhanced Definition Television，EDTV） EDTV为能够显示最高为576逐行有效扫描线的内容或显示器。没有定义宽高比。

均衡脉冲（Equalization Pulse） 这里有两组脉冲，一组发生在锯齿场同步之前，而另一组则发生在之后。这些脉冲的出现频率为正常水平扫描频率的两倍。它们的存在是为了确保在早期的电视中能够正确地实现2：1交替（interlacing）。

错误隐藏（Error Concealment） 能够隐藏传输错误的能力，这些错误在接收设备的能力之外对内容造成破坏，错误掩盖就是要使得能够对这种错误进行处理并以合适的方式显示内容。对于视频来讲，这些技术包括用各种数据来替换受损区域，这些数据可能是前面的视频数据、由前一帧和下一帧得到的插值视频数据，或当前帧内的邻居区域插值得到的视频数据。

错误恢复（Error Resilience） 能够处理传输错误的能力，并且不会由于超出接收设备的能力而对内容造成破坏，这样就可以用合适的方式对视频内容进行显示。MPEG-4支持通过使用重同步标记器（resynchronization marker）、扩展头编码、数据划分和可逆VLC进行错误恢复。

ETSI EN 300 163 该规范为PAL定义了NICAM 728数字音频。

ETSI EN 300 231 该规范定义了在场消隐期内发送的信息，在欧洲使用PAL图文电视（ETSI EN 300 706）来控制VCR。

ETSI EN 300 294 该规范为576i视频信号定义了宽屏信令（WSS）信息。对于576i视频系统，WSS可以在线23上提供。

ETSI EN 300 421 这是一个DVB-S规范。

ETSI EN 300 429 这是一个DVB-C规范。

ETSI EN 300 468 这是一个DVB SI（服务信息）规范。

ETSI EN 300 472 这是一个关于在DVB比特流上加载图文电视数据（ETSI EI 300 706）的规范。

ETSI EN 300 706 这是一个增强的PAL图文电视规范。

ETSI EN 300 708 该规范定义了如何使用PAL图文电视（ETSI EI 300 706）进行数据传输。

ETSI EN 300 743 这是一个DVB字幕规范。

ETSI EN 300 744 这是一个DVB-T规范。

ETSI EN 301 192 这是一个DVB数据广播规范。

ETSI EN 301 775 这是一个关于在DVB位流上承载场消隐信息（VBI）的规范。

ETSI EN 302 304 这是一个为手持市场定义的DVB-H规范。

ETSI EN 302 307 这是一个DVB-S2规范。

ETSI ETS 300 731 该规范定义了PAL＋（PALplus）标准，该标准准许在通常的PAL传输系统中传输16：9的节目。

ETSI ETS 300 732 该规范为PAL定义了重影消除参考（Ghost Cancellation Reference，GVR）信号。

淡入淡出（Fade） 淡入淡出是从一个视频源切换到另一个视频源的一种方法。以后，在看电视节目（或电影）的时候，当场景快要结束并转换到另一个场景时，特别注意的话就能看到图像的淡入淡出。场景淡出成为黑色，然后，又从黑色淡出为另一个场景。当场景之间的淡入淡出不需要达到黑色时，就称为淡化（dissolve）。实现淡入淡出的一种方式就是Alpha混合。

场（Field） 隔行显示是通过使用两个场形成的，每个场包含一半的扫描线，两个场共同构成了一个完整的视频帧。每个场都是完整地进行显示，所以，显示奇数场，接着是偶数场，然后又是奇数场，并一直持续下去。场并不仅仅在隔行扫描系统中存在。因此，对于480i的视频系统，其每帧数据有525线，场有262.5线，两个场组成一个525线的帧。

闪烁（Flicker） 当视频的帧率太低时就会出现闪烁现象。荧光照明灯具可以产生同样的效果。闪烁导致的两个问题是画面移动以及令眼睛感到疲劳。

FM 参见频率调制（Frequency Modulation）。

帧（Frame） 视频帧本质上讲是一幅图像，或仍在视频流之外的图像。足够快地播放这些单独的连续视频帧，看起来就像人在屏幕上运动一样。这与活页片、卡通动画和电影的原理是一样的。

帧缓冲（Frame Buffer） 帧缓冲用于存放一幅用于显示的图片的内存。我们得谈一下需要多少内存？好，假设水平分辨率为640个像素和480行扫描线，那么我们使用R'G'B'颜色空间。可以计算得到：

$$640 \times 480 \times 3 = 921600\text{B（即900KB）}$$

因此，在该分辨率下，需要900KB的内存来存储一帧视频数据。

帧率（Frame Rate） 视频源的帧率是指可以以什么样的速度得到新的静止图像。576i和480i显示器起初分别使用每秒25帧和每秒30帧的帧率。现在的刷新频率普遍是50帧/秒、60帧/秒、100帧/秒和120帧/秒。

帧率转换（Frame Rate Conversion） 帧率转换是将一种帧率转换为另一种帧率的行为。

脉冲前肩（Front Porch） 指的是模拟视频波形的一个区域，它位于行消隐和线同步的起始处。

FVD 终极通用磁盘（Finalized Versatile Disc）是一种增加容量的红光DVD磁盘，播放器规范来自中国台湾地区。它使用WMV9和WMA9编解码器。5.4GB/9.8GB的FVD-1磁盘支持多达135分钟的WMV9 720p的视频内容。6GB/11GB的FVD-2磁盘支持的视频内容可以达到1080i。

Gamma 大多数摄像机和显示器的传输特性（transfer characteristic）为非线性的。对于显示器来说，当信号电平较小时，幅度的微小变化将引起显示器亮度上的变化，但在信号电平较高时，同样的幅度变化不会引起同样的亮度变化幅度。这种非线性性就是我们所知的gamma特性。

Gamma校正（Gamma Correction） 在进行显示之前，必须对线性RGB数据进行处理（gamma校正）以补偿显示器的非线性性。

GCR 参见重影消除参考（Ghost Cancellation Reference）信号。

同步锁相（Genlock） 模拟视频信号为视频解码器提供所有必要的信息，以使得解码器可以重建图像。这包括亮度、颜色和时序信息。为了正确地对视频信号进行解码，视频解码器必须锁定所有嵌入到视频信号内的时序信息，包括色同步、线同步和场同步。解码器查看视频信号的色同步，并重建编码器使用的原始颜色副载波。为了能够正确地对颜色信息进行解码，副载波重建是需要的。它也生成采样时钟（通过查看视频信号内的同步信息来完成的），该时钟用于将来自解码器的像素数据送入内存或其他电路以便进一步处理。在解码器中完成所有这些工作的电路构成了同步锁相电路。尽管看起来简单，但是同步锁相电路必须能够处理很差的视频源，像VCR、相机和玩具摄像头的输出。实际上，同步锁相电路是视频解码器中最复杂的部分。

重影消除参考（Ghost Cancellation Reference） 位于（M）NTSC扫描线19和282上以及（B、D、G、H、I）PAL扫描线318上的参考信号，通过重影消除参考信号就可以消除TV中的鬼影。通过滤波来处理传输的GCR信号，并决定如何过滤整个视频信号以去掉鬼影。ITU-R BT.1124和ETSI ETS 300 732定义了用于每个国家的重影消除参考标准。ATSC A/49也定义了NTSC的重影消除参考标准。

灰度（Gray Scale） 术语灰度有多种意思。在某些情况下，它的意思是彩色视频信号的亮度分量。而在有的情况下，它又表示黑白视频信号。

H.261、H.263 建立ITU H.261和H.263视频压缩标准是为了在ISDN、LAN、普通电话线等上实现视频会议。H.261支持的视频分辨率为352×288和176×144，帧率可达29.97帧/秒。H.263

支持的分辨率为1048×1152、704×576、352×288、176×144和128×96，帧率可达29.97帧/秒。

H.264　“下一代”视频编解码。就是以前所知的H.26L、JVT和AVC（高级视频编解码），也就是现在的MPEG-4.10标准。

对于相同质量的视频，ITU-T H.264提供的最高比特率比MPEG-4高级简单档次（ASP）视频编解码比特率的50%还小。设计H.264是为了与SMPTE 421M（VC-1）视频编解码在比特率和质量上形成竞争。

H.26L　H.264视频编解码早期使用的名称。

HD-CIF　参见CIF。

HD-SDTI　高数据传输率串行数据传输接口，由SMPTE 348M定义。

HDMI　高分辨率多媒体接口（High-Definition Multimedia Interface）的缩写，这是一个用于消费类设备的单线数字音/视频接口。设计HDMI是为了以向后兼容的模式替换DVI，并支持CEA-861和HDCP。

支持传输率达5Gbit/s的数字RGB或YCbCr数据（HDTV要求2.2Gbit/s的数据传输率）。也可以支持多达8个32~192kHz数字音频通道，还支持AV链（AV.link）能力和更小的15mm 19针的连接头。

HDTV　参见高清电视（High-Definition Television）。

HDV　高清晰度数字视频（High-Definition Digital Video）的缩写，是一种高清晰度数字便携式摄像机格式。

高清电视（High-Definition Television，HDTV）　HDTV具有至少720逐行或1080隔行有效扫描线的显示能力。它必须能够以至少540逐行或810隔行有效扫描线显示16：9的图像。

高通滤波器（Highpass Filter）　一种电路，它所能通过的频率为特定频率（截止频率）以上的值。频率在截止频率以下的信号，其幅度被极大地削弱，这样就可以去掉这些信号。

行消隐（Horizontal Blanking）　在行消隐期间，当电子束在CRT显示屏上从右边到左边回扫时，等于消隐电平的视频信号并不显示电子束。

水平分辨率（Horizontal Resolution）　参见分辨率（Resolution）。

水平扫描速率（Horizontal Scan Rate）　它表示扫描电子束在显示器上从一边到另一边进行扫描的快慢。在480i视频系统中，该速率为63.556ms（即15.734kHz）。这意味着，每秒钟扫描电子束从一边移动到另一边共有15734次。

线同步（Horizontal Sync）　这是视频信号的一部分，它告诉显示器在从左到右的方向上什么地方放置图像。线同步脉冲告诉接收系统新的扫描线从什么地方开始。

House Sync　这是黑场信号（Black Burst）的另外一个名称。

HSYNC　请参考线同步（Horizontal Sync）的定义。

色调（Hue）　从技术的角度来讲，术语色度是指颜色的波长。这意味着，色调是用于基准色（红色、绿色和蓝色等）的一个术语。色调与颜色的强度和饱和度是完全独立的。例如，红色色调在低饱和度下看起来是棕色，在更高的饱和度上看起来是鲜红色，而在高亮度下看起来则是粉红色。所有这三种“颜色”具有相同的色调。

霍夫曼编码（Huffman Coding）　霍夫曼编码是一种数据压缩方法。它不关心数据表示什么含义（可以是图像数据、音频数据或其他的数据）。正因为如此，霍夫曼编码是JPEG、MPEG、H.261和H.263所采用的技术之一，使用霍夫曼编码有助于数据的压缩。下面就是其工作原理。首先观察一下需要压缩的数据，并创建一个表列出每片特定数据出现的次数。现在为出现得最频繁的那片数据指派一个非常短的码字。紧接着的下一个较大的码字指派给出现频率第二高的那片数

据。一直持续下去，直到每一个唯一的数据片都赋予了唯一长度可变的码字为止。其思想是，出现得最频繁的数据赋予一个小的码字，而很少出现的数据则赋予一个长的码字，这样就节省了空间。

HVD 高分辨率通用磁盘（High-definition Versatile Disc，HVD）是一种红光DVD磁盘，播放器规范来自中国。它支持可达150分钟720p的MPEG-2视频内容。播放器输出支持1080i和720p视频。

IDTV 参见改进清晰度电视机（Improved Definition Television）。

IEC 60461 为480i核576i视频系统定义了纵向时间码（LTC）和场消隐期时间码（VITC）。LTC需要在整个场时间（entire field time）内使用单独的途径来传输时间码（timecode）信息。VITC在场消隐期内每个场使用一行扫描线。也可以参见SMPTE 12M。

IEC 60958 为消费类（SPDIF）和专业类应用定义了串行数字音频接口。

IEC 61834 定义了DV［起初为“蓝皮书（Blue Book）”］标准。也可以参见BT.1618和SMPTE 314M。

IEC 61880 为480i视频信号定义了宽屏信令（WSS）信息。可以在线20和283上提供WSS。

IEC 61883 定义了在IEEE 1394上传输数据、音频、DV（IEC 61834）和MPEG-2的方法。

IEC 62107 定义了超级VCD标准。

IEEE 1394 一种高速菊链（daisy-chained）串行接口。数字音频、视频和数据可以在确保带宽或确保延迟条件下传输。它不是可热插拔的，使用小的6针或4针连接头，具有6针的连接头可以供电。

非法视频（Illegal Video） 在NTSC和PAL视频领域中，在R'G'B'颜色空间中有些颜色是不能表示的。例如，由于颜色空间的限制，在R'G'B'颜色空间中100%饱和度的红色（即完全强度的红色，没有绿色和蓝色值）在NTSC视频信号中是不存在的。NTSC编码器必须能够确定正在生成的颜色是无效颜色（illegal color），并停止生成这样的颜色，因为这可能会导致过饱和（oversaturation）和溢处（blooming）。

改进清晰度电视机（Improved Definition Television，IDTV） IDTV不同于HDTV。IDTV是一个系统，它通过在电视机中增加一些处理来改善电视机上的显示效果，发送的信号是标准的NTSC或PAL信号。

强度（Intensity） 它与亮度（lightness）表示同样的意思。

隔行（Interlaced） 隔行视频系统是这样一种系统，它是由两个交替的场（field）来生成一个完整的视频帧。所以，一个场中的线数为一帧视频总线数的一半。在480i视频系统中，每个场有262.5线（一帧有525行扫描线），而在576i视频系统中，每个场有312.5线（一帧有625线）。每个场连续地在显示屏上绘制——首先绘制一个场，然后另外一个场。

插值（Interpolation） 插值是生成额外信息的一种数学方式。假如说，一幅图像需要放大一倍，即从100个样本放大到200个样本。“缺失”的样本就需要根据现有的样本通过计算得到的新样本来补充。在计算好所有的新样本以后（瞬间完成），在原来只有100个样本的地方现在有200个样本了，并且现在的图像大小是原来的两倍了。

IRU单位（IRE Unit） 一种用于描述视频信号幅度特性的任意单位（arbitrary unit）。白色定义为100IRU，消隐电平定义为0IRU。

ISDB 综合服务数字广播（Integrated Service Digital Broadcasting）的缩写，日本使用的数字电视（DTV）广播标准。

ITU-R BT.xxx 参见BT.xxx。

抖动（Jitter） 信号特征（如频率、幅度等）的短期波动。

JPEG JPEG表示静止图像专家组（Joint Photographic Experts Group）。然而，当人们在使用术语JPEG时，其含义通常是由他们开发的图像压缩标准。开发JPEG是为了对静止图像进行压缩，如照片、单个视频帧、扫描到计算机中的图像等。可以根据应用所要求的任意速度运行JPEG。对于静止图像数据库，算法并不一定非常快。如果JPEG足够快，就可以压缩运动的视频，这意味着JPEG将必须以每秒50场或60场的速度运行。这就是所谓的运动JPEG或M-JPEG。如果设计者正在设计一个视频编辑系统，那么就有可能想这么干。现在，运行速度为每秒60场的M-JPEG并不比每秒60场的MPEG-2高效，这是因为MPEG的设计利用了运动视频的某些特征。

JVT 视频联合工作组（Joint Video Team）的缩写，一群主要致力于新视频编解码研究的视频专家。它是国际电信联盟（ITU）、国际电工委员会（IEC）和国际标准化组织（ISO）之间共同协作的结果。JVT也是H.264视频编解码早期使用的名称。

kbit/s 千比特/秒（kilobits per second）的缩写。

KB/s 千字节/秒（kilobytes per second）的缩写。

行锁时钟（Line-Locked Clock） 一种设计，它保证了每行扫描线总是有固定数目的样本，即使行时序发生了变化也是如此。

行存储（Line Store） 行存储是一种内存，是用来保存一行视频数据的。如果有效显示器的水平分辨率为640个样本，以R′G′B′作为颜色空间，那么行存储将必须具有640个样本位置（location）的长度，宽度为3B。这相当于一个样本的一个颜色占据一个位置。典型情况下，滤波算法需要用到行存储。例如，梳状滤波器就是由一个或多个行存储组成的。

线性性（Linearity） 线性性衡量ADC或DAC变换好坏的一个基本测度。典型情况下，线性性是通过使ADC或DAC试图生成一个线性增加的信号。将实际输出与理想输出进行比较。二者之间的差异就是线性性的测度。这个数越小就越好。典型情况下，线性性定义为一个区间或最低有效位（LSB）的百分比。

锁定（Locked） 当PLL正在准确地产生时序、并且这些时序与输入视频源的时序是对齐的时，就说PLL是锁定的。当PLL是锁定的时，PLL就处于稳定状态，在生成的采样时钟中就只存在最小的扰动。

纵向时间码（Longitudinal Timecode，LTC） 时间码信息存在于视频的独立通道中，要求在整个场时间内保存和读取时间码。

无损的（Lossless） 无损是用于压缩领域的一个术语。当压缩后的数据还是与原始数据完全一样时，就称压缩是无损压缩。之所以是无损的，是因为没有丢失任何信息。

有损的（Lossy） 有损压缩是与无损压缩刚好相反的。从压缩数据中重新生成的数据与原始数据是不一样的。这种差异既可能是值得我们注意的，也可能是无关紧要的。但只要压缩前后的两个图像是不一样的，那么压缩就是有损的。

低通滤波器（Lowass Filter） 是一种电路，能通过的频率为特定频率（截止频率）以下的频率。截止频率以上的频率在幅度上被减小，其目的只是为了去掉高频部分的信号。

LTC 参见纵向时间码（Longitudinal Timecode）。

亮度（Luma） 与上面提到的色度定义一样，视频系统使用的信号具有两个部分：黑白部分和彩色部分。黑白部分为亮度。正是亮度分量才准许彩色电视广播可以被黑白电视所接收，并仍然保持可以观看。

亮度（Luminance） 在视频中，术语亮度（Luminance）和亮度（Luma）通常是可交换

使用的。参见亮度（Luma）的定义。

Mbit/s 兆比特/秒的简写。

MB/s 兆字节/秒的简写。

MESECAM 记录SECAM视频的一种技术。MESECAM技术不是将FM颜色副载波除以4，然后在播放时进行相乘以还原原来的副载波，而是采用与PAL相同的外差变换（heterodyne conversion）。

MHP 参见多媒体家庭平台（Multimedia Home Platform）。

M-JPEG 参见运动JPEG（Motion JPEG）。

调制器（Modulator） 从根本上讲，调制器是一种电路，它将两种信号结合起来，其结合方式是要使得以后可以将它们分开。这与视频有什么关系呢？让我们以NTSC系统为例，当然这个例子应用到PAL中也非常好。NTSC系统可能会使用YIQ或YUV颜色空间，并且I和Q或U和V信号包含了图像中的所有颜色信息。在I和Q或U和V分量上调制了两个3.58MHz的颜色副载波（相位相差90度），这两个副载波加在一起就可以建立NTSC视频的色度部分。

波纹（Moire） 这是一种类型的图像伪影。当创建一个显示器上实际不应该有的模式时，就会出现波纹效应。典型情况下，当两个不同的频率发生作用并产生不希望的新频率时，就会产生波纹模式。

单色（Monochrome） 单色信号是一种视频源，它只有一个分量。尽管通常是指亮度（或黑白）视频信号，但红色视频信号也是单色，因为它也仅有一个分量。

单调的（Monotonic） 这是用于描述ADC和DAC的一个术语。如果对于在输入信号的增加都对应地有一个输出信号的增加，那么就说ADC或DAC是单调的。任何单调的DAC或ADC都没有实际作用（这意味着输出随着输入的减小而减小），没有人会用一个单调的ADC或DAC。

运动估计（Motion Estimation） 运动估计就是要尽可能地描述目标从一个视频帧运动到另一个视频帧的何处。那么为什么要估计目标的位置呢？好，让我们举一个视频源的例子。在该视频源中，一个球从空中飞过。背景是单色，并且不同于球的颜色。球在一个视频帧中处于某个位置，而在下一个视频帧中则已向上运动了，并偏向右方。现在假设视频摄像机仅仅发送出了该视频序列的第一个视频帧，不再发送第二个视频帧而仅仅发送球的位置是不是更有效呢？除了球之外没有别的东西在运动，因此仅需要发送两个小的数据。这就是运动估计的本质所在。顺便说一下，运动估计是MPEG、H.261和H.263的完整组成部分。

运动JPEG（Motion JPEG） 用于对视频进行实时压缩与解压的JPEG。视频的每个场或帧是独立进行处理的。

MPEG 运动图像专家组（Moving Picture Experts Group）。这是一个ISO/IEC（国际标准化组织）体，它正致力于各种压缩算法的开发。MPEG与JPEG的不同之处在于，MPEG利用了运动图像序列中帧到帧之间的信息冗余，而JPEG则没有。

MPEG-1 第一个MPEG标准，定义了实时音视频的压缩格式。尽管也可以支持更高的分辨率，但典型情况下的视频分辨率为352×240或352×288。最大的比特率大约为1.5Mbit/s。VCD采用MPEG-1格式。

MPEG-2 MPEG-2对MPEG-1标准进行了扩展以满足更广范围内的应用。支持更高的视频分辨率，这样就可以满足HDTV应用，并同时支持逐行和隔行视频。

MPEG-3 起初MPEG-3的目标是HDTV应用。由于HDTV应用已经包含到了MPEG-2中，因此也就没有MPEG-3标准了。

MPEG-4 MPEG-4（ISO/IEC 14496）支持面向对象的方法，其中，场景被建模为对象的

组合，这些对象可以是自然对象，也可以是合成对象，用户可以与这些对象进行交互。场景中的视频对象可以在数学上进行描述，也可以在二维或三维空间中指定其位置。类似地，音频对象可以放在声音空间中。这样，仅需要定义一次音频对象和视频对象即可，观众可以改变自身的观看位置，在局部进行音视频更新的计算。经典的矩形视频，如来自相机的视频，这是一种受支持的可视对象。另外，可以将图像映射到计算机生成的各种形状以及文语转换（text-to-speech）接口上。

MPEG-4标准已经将H.264（下一代视频编解码）作为第10部分包含进来了。

MTS 多频道电视声音（Multichannel Television Sound）的缩写。为各种立体声音频实现的一般名称，如BTSC和Zweiton。

NABTS 北美广播图文电视规范（North American Broadcast Teletext Specification）EIA-516的缩写。这也是ITU-R BT.653 525线系统C图文电视。然而，NABTS规范要详细得多。

NexTView 基于ETSI ETS 300 707的电子节目指南（Electronic Program Guide，EPG）。

NICAM 728 一种使用另一个音频副载波为PAL视频实现数字立体声的技术。比特率为728kbit/s。相关的讨论见BS.707和ETSI EN 300 163。在中国，NICAM 728也用于传送非音频数字数据。

非隔行（Noninterlaced） 这是一种对显示屏进行扫描的方法，它是与隔行方法完全相反的一种扫描方法。一帧中的所有线顺序地进行扫描，一行紧接着一行的进行。术语“场”并不适用于非隔行系统。非隔行系统中的另一个术语是逐行扫描。

NTSC 它可能是永无重色（Never Twice the Same Color 或Never The Same Color）或国际电视标准委员会（National Television Standard Committee），这取决于上下文。在技术上，NTSC只是一种颜色调制策略而已。为了完全规定彩色视频信号，应该称之为（M）NTSC。NTSC通常（尽管不正确）也用于指525/59.94或525/60中的任一种视频系统。参见NTSC 4.43。

NTSC 4.43 一种NTSC视频信号，它使用的是PAL颜色副载波频率（大约为4.43MHz）。是由索尼公司在1970年开发的，是为了方便地用欧洲接收设备来接收NTSC信号。

NVOD 准视频点播（Near-Video-On-Demand）的缩写。参见视频点播（Video-On-Demand）。

OIRT 国际广播电视组织（Organisation Internationale de Radiodiffusion et Television）的缩写。

开放字幕（Open Subtitle） 参见字幕（Subtitle）。

过采样VBI数据（Oversampled VBI Data） 参见原始VBI数据（Raw VBI Data）。

过扫描（Overscan） 当显示一副图像时，如果有一小部分图像超出了显示屏的边界，那么该图像就是过扫描的。过扫描在所有的电视中都是很常见的。

PAL PAL可以表示逐行倒相彩色电视制式（Phase Alternation Line）、画面总是糟糕（Picture Always Lousy）或终极完美（Perfect At Last），这取决于具体环境。在技术上，PAL只是一种颜色调制策略。为了完全指定彩色视频信号，应该称之为（B、D、G、H、I、M、N或NC）PAL。（B、D、G、H、I）PAL为用于欧洲和其他很多国家的彩色视频标准。PAL通常（尽管不正确）也用于指任意的625/50视频系统。也可以参见PAL 60。

PAL 60 一种NTSC视频信号，它使用的是PAL颜色副载波频率（大约为4.43MHz）和PAL类型的颜色调制。它更进一步地适应了NTSC 4.43，除改变了颜色副载波频率外，还修改了颜色调制方式。它是由JVC于上世纪80年代开发的，这也是为了配合他们的视频盘播放器，因此其早期的名字为“磁盘PAL（Disk-PAL）”。

有一个很少使用的变种，也称为PAL 60，它是一种PAL视频信号，使用的是NTSC颜色副载波频率（大约为3.58MHz）和PAL类型的颜色调制方式。

PAL＋（PALplus） PAL＋是16：9宽高比版本的PAL，设计它是为了使用普通的PAL进行传输。没有PAL＋解码器的16：9的电视机和标准的4：3电视机都显示标准画面。它是在BT.1197和ETSI ETS 300 731中定义的。

PDC 参见节目传输控制（Program Delivery Control）。

基底电平（Pedestal） 基底电平[①]就是一种偏差，用于通过一个小的差值将黑色电平与消隐电平区分开。当视频系统不使用基底电平时，黑色和消隐电平就是一样的了。

相位调整（Phase Adjust） 这是一个用于在NTSC视频信号中描述色度调整方法的术语。相对于颜色电平移，颜色副载波的相位发生移动，或进行了调整。PAL和SECAM系统通常没有相位（或色度）调整控制。

像素（Pixel） 像素是图元（picture element）的简称，像素为组成扫描线的最小样本。例如，当水平分辨率定义为640像素时，这意味着有640个独立的位置或样本，它们组成了每行扫描线的可见部分。像素可以是正方形或长方形。

像素时钟（Pixel Clock） 像素时钟用于将视频的水平线划分为样本。像素时钟相对于视频必须保持稳定（非常小的波动），否则图像就不能正确地保存。像素时钟的频率越高，每线就有越多的样本。

像素失落（Pixel Drop-Out） 这是一个真正的麻烦制造者，因为它可能引起伪影。有些情况下，在屏幕上，像素失落看起来像黑点，既可能是静止的，也可能是在四周移动。有几种可能会引起像素失落，像ADC没正确地对视频进行数字化。ADC和帧缓冲之间的时序不正确，这就导致内存中存储了错误的数值，这也可能引起像素失落。因此，在视频系统的任何地方，时序都可能引起像素失落。

基准色（Primary Colors） 为一组颜色，通过它们的结合可以产生任意所需的中间颜色集，在边界之内的这些颜色称为“色域（gamut）”。彩色电视的基准色为红、绿和蓝。使用的准确红色、绿色和蓝色取决于电视标准。显示器设备通常不使用相同的基准色，这导致与理想情况有微小的出入。

节目传输控制（Program Delivery Control） 在欧洲，使用图文电视在场消隐期内发送信息来控制VCR。对应的规范为ETSI ETS 300 231。

逐行扫描（Progressive Scan） 参见非隔行扫描（Noninterlaced）。

伪色（Pseudo Color） 伪色是用于描述一种技术的术语，该技术就是应用或显示实际上不存在的颜色。对这样一类卫星图像我们都很熟悉，就是在晚间天气预报时，显示的跨省温差图，或多种颜色的云运动形成的序列图。这些都是现实中伪色的例子。这些颜色并非真实存在。计算机使用查询表内存来增加这些颜色，对像温度或云高度这样的数据进行可视化。

Px64 该术语基本上与H.261是一样的。它已经不再使用了，因为H.261被用在ISDN视频会议以外的应用中了。

QAM 参见正交调幅（Quadrature Amplitude Modulation）。

QCIF 四分之一通用接口格式（Quarter Common Interface Format）。该视频格式是考虑用于实现低价的视频电话。QCIF格式的分辨率为176×144的有效像素，帧率为每秒29.97帧。

QSIF 四分之一标准接口格式（Quarter Standard Interface Format）。使用方形像素的计算机

① 黑色电平与消隐电平之差。——译者注

公司，将ASIF定义为160×120有效像素，帧率可以为计算机可以支持的任意帧率。

四倍色度（Quad Chroma） 四倍色度是指一种技术，即采样时钟频率为色同步频率的四倍。对于NTSC来讲，这意味着，采样时钟大约为14.32MHz（4×3.579 545MHz），而对于PAL来说，采样时钟则大约为17.73MHz（4×4.433 618 75MHz）。之所以这些频率成为普遍使用的采样时钟频率，是因为根据所选择的方法，它们可以使得色度（颜色）的编解码更加容易。

正交调幅（Quadrature Amplitude Modulation） 这是将数字数据编码到载波上用于射频发射的一种方法。典型情况下，QAM用于有线电视以传输DTV信号。DVB-C支持QAM、32-QAM、64-QAM、128-QAM和256-QAM，但接收设备仅需要支持64-QAM以下的各种调制。

正交调制（Quadrature Modulation） 这是将相位距离为90度的两个载波分量调制在一起的一种调制方法。

量化（Quantization） 将连续模拟信号转换为一组离散电平（数字化）的过程。

量化噪声（Quantization Noise） 这是一种在量化过程中引入的一种固有的不确定性，这是由于量化以后只产生了离散的、而不是连续的电平。量化噪声也称为量化失真。

光栅化（Raster） 本质上讲，光栅化是组成一幅画面的一组扫描线。有人可能会经常听到术语光栅线（raster line），它与扫描线是同一个东西。组成一个视频帧的所有扫描线形成了一个光栅。

原始VBI数据（Raw VBI Data） 这是一种技术，在这种技术中，VBI数据（如图文电视和字幕数据）由快速的采样时钟进行采样（比如27MHz）和输出。这种技术准许通过软件方式完成VBI数据的解码工作。

RC时间码（RC Timecode） 可重写的时间码，用于消费类视频产品。

实时控制协议（Real Time Control Protocol） 参见RTCP。

实时流化协议（Real Time Streaming Protocol） 参见RTSP。

实时传输协议（Real Time Transport Protocol） 参见RTP。

矩形像素（Rectangular Pixels） 不是方形像素就是矩形像素。

残留副载波（Residual Subcarrier） 这是指复合视频信号白色、灰色或黑色区域中提供的颜色副载波的信息量（理想情况下，应该是没有信息的）。这个数通常表示为$-n$ dB，n越大越好。

分辨率（Resolution） 这是衡量一幅图像中有多少信息是可见的一种基本测度。通常描述为$h \times v$。h为水平分辨率（横跨显示器），v为垂直分辨率（沿显示器的上下方向）。这两个数越大越好，因为，这意味着可以看到的细节更多。如果仅指定了一个数，那么就应该是水平分辨率。

显示器规定了它所能处理的最大分辨率，这是由显示器技术和所使用的电子设备所决定的。实际的分辨率为视频源和显示器中较低的分辨率。

垂直分辨率为一幅画面从上到下可见的白色到黑色和黑色到白色的变换次数。最大数目为图像所使用的有效扫描线数目。实际的垂直分辨率可能比这要少，这是由于处理、隔行、过扫描或视频源的限制等原因所导致的。

水平分辨率为一幅画面上从左到右可见的白色到黑色和黑色到白色的变换次数。对于数字显示器而言，最大数目为扫描线使用的有效像素的数目。对于模拟和数字显示器而言，实际的垂直分辨率可能比这要少，这是由于处理、过扫描或视频源的限制等原因所导致的。

资源预留协议（Resource Reservation Protocol） 参见RSVP。

回扫（Retrace） 回扫是指当电子束扫描到CRT显示器的右边缘，然后回到左边缘的过

程。回扫是在水平消隐期内进行的。

RGB 为红色、绿色和蓝色的简写。RGB用于表示线性RGB数据。R′G′B′用于表示经过Gamma校正的RGB数据。

RS-170、RS-170A RS-170为美国标准，被用于黑白电视，并定义了电压电平、消隐时间、同步脉冲宽度等。该规范给出了接收设备用于显示单色图像所需的所有定义。现在，SMPTE 170M本质上与该规范是一样的，只是通过增加颜色分量进行了改进，使得其适合于彩色电视。只是对RS-170进行了一点细微的修改，这是为了可以增加颜色（RS-170A），最终结果成为NTSC的SMPTE 170M标准。这个微小的改变是如此之小，以至于现有的黑白电视甚至不会注意到这种变化。

RS-343 RS-343与RS-170完成的工作是同样的，定义了变换模拟视频的规范，但差异为RS-343面向的是高分辨率计算机图形模拟视频，而RS-170 则面向TV分辨率的NTSC模拟视频。

RSVP RSVP（资源预留协议）为一个控制协议，它准许接收设备在IP网络上申请特定的服务质量水平。像视频流这样的实时应用使用RSVP协议在传输路径的路由器上预留必要的资源，这样当传输实际进行时，请求的带宽就可以得到保证。

RTCP RTCP（实时控制协议）是一种控制协议，是用于和RTP协同使用的。在RTP会话期间，参与会话的各方周期性地发送RTCP包以传输服务质量和成员管理相关的状态信息。RTCP也使用RSVP协议来预留资源，以保证特定的服务质量得以满足。

RTP RTP（实时传输协议）是一种包格式，也是用于在IP网络上传输实时音视频数据的协议。这些实时数据可以是任意的格式，包括MPEG-2、MPEG-4、ASP、QuickTime等。实现时间重建、实效检测、安全和内容鉴别，它还支持实时音视频的多播（从一个源发送到多个目的地）和单播（从一个源发送到一个目的地）。支持一路传输（如视频点播）和交互式服务（如网络电话）。RTP是用于和RTPC协同工作的。

RTSP RTSP（实时流化协议）是一个客户－服务器协议，它使得可以在IP网络上实现音视频流的可控传输。它提供了VCR模式的远程控制能力，如播放、暂停、快进和倒带。实际的数据传输是通过RTP协议完成的。

游程编码（Run-Length Coding） 游程编码是一种数据压缩类型。比如说，这个页面足够宽，它每行可以容纳80个字符。现在，想象一下，有一行除了几个字以外，都是空格。尽管它有80个字符长，但是几乎都是空格，比如说，在词coding和medium之间有50个空格。这50个空格需要存储为50个单独的编码，但这需要占掉50字节的存储量。一种替代方式是，定义一个特定的编码表示接下来是一串空格，接着再用一个数字表示空格串中空格的个数。因此，对于上面我们举的例子而言，仅需要两个字节来表示50个空格，第一个字节表示特定的编码类型，接着的一个字节表示空格数50。我们压缩了数据，将50字节压缩为2字节。其压缩率为25：1。这还不错，除了仅仅对一个文档中的一行数据进行压缩外，如果对整个文档进行压缩，总的压缩率应该比这小得多。

如果仅仅使用游程编码对图像进行压缩，其效率是不如DCT的压缩效率的，因为在真正的图像中很少存在长度相同的长行程。游程编码相对于DCT编码而言，其仅有的优点就是易于实现。尽管游程编码本身对于图像的压缩效率不高，但是它仍然作为其中的一部分用于JPEG、MPEG、H.261和H.263压缩方案。

R-Y 在视频中，为红色－亮度信号，也称为色差信号，当它加上亮度（L）信号时，就得到了红色信号。

采样/样本（Sample） 获取周期区间处的信号值。也表示在给定时刻处信号的值。

采样与保持（Sample and Hold） 一种电路，它对信号进行采样并保持采样值，直到开始进行下一个采样为止。

采样率（Sample Rate） 采样率是对信号进行采样的速度。采样率由采样时钟频率决定。

SAP 参见辅助音频节目（Secondary Audio Program）。

Saturation（**饱和度**） 为颜色提供的色彩数量（amount of color）。例如，饱和度较小的红色看起来是粉红色，而满饱和度的红色看起来则是红蜡笔的颜色。饱和度并不是颜色的亮度，仅表示一个颜色中有多少色素。色素越少，颜色就越不饱和，其效果就像在纯色中增加了白色一样。

SBTVD-T 由巴西采用的地面数字电视广播系统。射频调制与ISDB-T是一样的。

缩放（Scaling） 缩放是改变图像分辨率的一种行为。例如，将一幅大小为640×480的图像缩小一半，其结果为一幅320×240的图像。放大一倍则得到一幅大小为1280×960的图像。对图像进行缩放有很多不同的方法，其中一些方法看起来要优于另外一些。但是，通常算法看起来越好，其实现代价也越高。

扫描线（Scan Line） 扫描线为横跨显示器的单独的线。525行这样的扫描线组成一幅480i或480p的电视画面，625行这样的扫描线组成一幅576i或576p的电视画面。

扫描速度调制（Scan Velocity Modulation） 参见速度扫描调制（Velocity Scan Modulation）。

SCART 一种欧洲加密电视接收机的接口（Syndicat des Constructeur d'Appareils Radio Recepteurs et Televiseurs）。它是一个21针的连接头，欧洲很多的消费类电子产品都支持该接口。它准许在设备之间传输单声道或立体声道音频，以及复合、S-Video或RGB视频。

IEC 60933-1和60933-2标准规定了基本的SCART连接头，包括信号电平。

scRGB scRGB颜色空间（正式称为sRGB64）对sRGB的动态范围、色域和位精度进行了扩展。scRGB色域不仅比sRGB的色域要大得多，而且比人类可以看到的颜色范围还要大。针对scRGB（IEC 61966-2-2）的规范使用BT.709的色度、D65参考白色和线性RGB数据（每种颜色为16位）。

不再使用0~1的归一化取值范围，所支持的取值范围为－0.5~＋7.4999。尽管二者具有相同的基准色，但在0之下和1之上的值就是scRGB比sRGB所多出的色域范围。

SDI 串行数字I/O。它是由BT.656定义的270Mbit/s或360Mbit/s串行接口的另一个名称。

SDTI 串行数据传输接口（Serial Data Transport Interface），由SMPTE 305M定义。

SDTV 参见标准清晰度电视（Standard-Definition Television）。

SECAM 这是类似于PAL的另一种彩色视频格式。二者之间主要差异在于SECAM的色度为FM调制，R-Y和B-Y信号按照线顺序进行传输。SECAM代表带有内存的顺彩（Sequentiel Couleur Ave Memoire或Sequential Color with Memory）。

辅助音频节目（Secondary Audio Program） 通常用于以第二种语言传输音频信号。也可以用于传输节目中关键视觉元素的[语音]描述，这些描述插入到节目音频的自然停顿中。

锯齿脉冲（Serration Pulses） 这是NTSC、PAL和SECAM中发生在场同步间隔期的一些脉冲，脉冲速率为普通水平扫描速率的两倍。存在锯齿脉冲的原因是为了确保在早期的电视中能够正确地实现2：1交替和消除直流分量偏移的积累。

基准电平（Setup） 与基准电平（Pedestal）是同一个东西。

SIF 标准（或源）输出格式。这种视频格式是为了实现数字视频的存储和传输。625/50 SIF格式的分辨率为352×288有效像素，帧率为每秒25帧。525/60 SIF格式的分辨率为352×240有

效像素，帧率为每秒30帧。注意，MPEG-1准许的分辨率可以达到4095×4095有效像素，但是，有一个定义为SIF的参数"约束子集"。使用方形像素的计算机企业将SIF定义为320×240有效像素，帧率为计算机所能支持的任意帧率。

信噪比（Signal-to-Noise Ratio，SNR） 信噪比是信号幅度与对信号形成干扰的不希望的信号（噪声）量之比。SNR通常定义为分贝（dB）。比值越大，图像看起来就越好。

无线电静默（Silent Radio） 无线电静默是一种服务，它提供的数据通常为宾馆和夜总会中看到的一些信息。通常为红色的大字，主要是显示当前的新闻、事件、积分等。这些信息在NTSC的线10~11和273~274上提供，使用类似于CEA-608的编码方式。

分片VBI数据（Sliced VBI Data） 这是一种技术，在这种技术中，VBI解码器对VBI数据（如图文电视和字幕数据）进行采样，锁定时序信息，并将其转换为二进制的0和1。必须对直流分量偏移、幅度波动和伪影进行补偿，这是为了通过VBI解码器可以准确地恢复数据。

SMPTE 12M 为480i和576i系统定义了纵向时间码（LTC）和场消隐期时间码（VITC）。LTC要求整个场时间内，使用独立的途径（track）存储时间码信息。VITC在场消隐期间，每个场中使用一行扫描线。

SMPTE 125M 720×480专业视频隔行标准（29.97Hz）。包括数字表示和数字并行接口。也可以参考BT.601和BT.656。

SMPTE 170M 美国使用的NTSC视频规范。参见RS-170A和BT.470。

SMPTE 240M 1920×1035专业视频隔行标准（29.97Hz或30Hz）。包括模拟RGB和YCbCr表示。这是由SMPTE 260M定义的数字并行接口。

SMPTE 244M 768×486专业视频隔行标准（29.97Hz）。包括数字表示（以4倍F_{SC}采样的复合NTSC视频）和数字并行接口。该数字串行接口是由SMPTE 259M定义的。

SMPTE 253M 用于专业视频SDTV系统的模拟RGB视频接口规范。

SMPTE 267M 960×480专业视频隔行标准（29.97MHz），包括数字表示和数字并行接口。也可以参考BT.601和BT.1302。

SMPTE 274M 1920×1080专业视频隔行和逐行标准（29.97、30、59.94和60MHz）。包括数字表示、模拟RGB和YPbPr接口和数字并行接口。数字串行接口是由SMPTE 292M定义的。

SMPTE 293M 720×480专业视频逐行标准（59.94 MHz）。包括数字表示、模拟RGB和YPbPr接口以及数字并行接口。数字串行接口是由SMPTE 294M定义的。也可以参考BT.1358和BT.1362。

SMPTE 296M 1280×720专业视频逐行标准。包括数字表示、模拟RGB和YPbPr接口。数字并行接口使用SMPTE 274M。数字串行接口是由SMPTE 292M定义的。

SMPTE 305M 串行数据传输接口（SDTI）。这是一种基于BT.656的270Mbit/s或360Mbit/s的串行接口，可以用于传输几乎任意类型的数字数据，包括MPEG-2节目和传输比特流、数字视频比特流等。不可以在使用不同数据类型的设备之间交换信息。以一种数据类型创建的信息仅可以传输到其他支持相同数据类型的设备。有独立的映射文档，它们描述了每种数据类型进入305M传输的格式。

SMPTE 314M 定义了DV视频、数据和压缩视频以25Mbit/s和50Mbit/s传输速率进行传输的数据结构。也可以参考BT.1618和IEC 61834。

SMPTE 344M 为专业视频应用定义了540Mbit/s串行数字接口。

SMPTE 348M 高数据率串行数据传输接口（HD-SDTI）。这是一个基于SMPTE 292M的1.485Gbit/s的串行接口，它可以用于传输几乎所有类型的数字数据，包括MPEG-2节目和传输位流、

DV位流等。不能在使用不用数据类型的设备之间交换信息。以一种数据格式定义的信息只能传输到支持相同数据结构的其他设备。有独立的映射文档描述每种数据类型进入348M传输的格式。

SMPTE 370M 这个SMPTE标准定义了数据率为100Mbit/s的基于DV的音频、数据和压缩视频。也可以参考BT.1620。

SMPTE 421M 这个SMPTE标准起初为我们所知的VC-1，它是基于微软的WMV9视频编解码的。设计它是为了与MPEG-4.10（H.264）进行竞争。

SMPTE RP160 用于专业视频HDTV系统的模拟RGB和YPbPr视频接口规范。

SPDIF 索尼/飞利浦数字接口（Sony/Philips Digital InterFace）的缩写。这是一个消费类接口，用于传输压缩或2通道LPCM数字音频。使用的是串行、自时钟（self-clocking）策略，它是基于同轴电缆或光纤连接的。现在IEC 60958完全定义了这个接口的消费类和专业应用。

分离同步加扰（Split Sync Scrambling） 分离同步是一种视频加扰技术，通常它是与行消隐翻转（inversion）、有效视频翻转之一或与二者同时使用。在分离同步技术中，将水平脉冲分为两半，其中脉冲的第二半的电平位于＋100IRE处，而不是位于标准的－40IRE处。根据不同的加扰模式，可能存在各种翻转情况：整个水平消隐期以＋30IRE为轴翻转，有效视频围绕＋30IRE为轴翻转，二者都翻转，或二者皆不翻转。通过将线同步脉冲分离，解扰器可以使用－40IRE和＋100IRE的参考电平。

由于部分线同步仍然处在－40IRE处，有的同步分离器仍然可以锁定缩短的线同步脉冲。当然，查询色同步的时序电路在线同步结束之后的固定间隔内可能发生混乱。另外，如果有效视频发生了翻转，有些视频信息可能会降到0IRE之下，这可能会扰乱同步监测电路。

总是以正确的频率和时序提供同步（burst），但是，当水平消隐期发生翻转后，相位就平移了180°。

方形像素（Square Pixels） 当每线的有效像素与每帧的有效线之比与显示器的宽高比相同时，这样的像素就是方形像素。这是与这样一个采样网格相同的，即该网格中的每个样本点的垂直空间与水平空间是相等的。

sRGB 针对sRGB（IEC 61966-2-1）的规范，使用BT.709色度、D65参考白色、显示器gamma值为2.2和线性RGB（每个颜色8位）。

sRGB值具有归一化的取值范围0~1，8位sRGB值，其黑白颜色值的取值范围为0~255。一个称为“演播室（Studio）RGB”的版本，使用表示范围为16~235的8位数表示黑白色，这是为了使得与视频应用兼容。

sRGB存在一个局限，由于归一化取值被限定于0~1，所以在这个色域之外的颜色（由它们生成的颜色三角形）不能使用。鉴于这个原因，开发了扩展的RGB颜色空间，即scRGB。

标准清晰度电视（Standard Definition Television，SDTV） SDTV指的是能够显示最多576行隔行有效扫描线的视频内容或显示器。没有定义宽高比。

流视频（Streaming Video） 用于在因特网或其他网络上实时传输的压缩音视频数据。它通常提供VCR模式的远程控制能力，如播放、暂停、快进和倒带。

副载波（Subcarrier） 一种辅助信号，它包含加入到主信号中的额外信息。

欠采样（Subsample） 欠采样表示以较低的速率进行采样的信号，而系统中某些其他的信号则不是。欠采样的主要例子为4：2：2 YCbCr颜色空间。对于每两个亮度（Y）样本，仅有一个Cb和一个Cr样本出现。这意味着Cb和Cr是欠采样的。

字幕（Subtitle） 加入到画面下方或画面之上的文本，其内容通常是与语音对应的文字，可能是与语音使用的语言不同的另外一种语言。开放字幕（open subtitle）作为已经加上字幕的视

频进行传输。封闭字幕（closed subtitle）在VBI期间传输，依赖于电视机对其解码，以及是放在画面上还是画面下方。隐藏字幕（closed captioning）是字幕（subtitling）的一种形式。ETSI ETS 300 743中定义了DVB的字幕。

超黑（Superblack） 嵌入到视频信号中的一种抠像信号，其电平位于黑色和同步电平之间。通常超黑信号是用于改善亮度自抠像（luma self-keying）的，因为视频信号包含黑色电平，这使得好的亮度自键难于实现。当顺向位移抠像器（downstream keyer）检测超黑电平时，它将插入另一个视频信号。

超级VCD（Super VideoCD，Super VCD，SVCD） 由中国国家技术标准委员会定义的关于记录方面的标准，该CD标准可以存放35~70分钟的数字音频和视频信息。使用的是MPEG-2视频，分辨率为480×480（29.97Hz的帧率）或480×576（25Hz的帧率）。音频使用MPEG层2（MPEG layer 2），比特率为32~384kbit/s，支持4个单通道、2个立体声或5.1通道。字幕使用覆盖而不是子画面（DVD视频）或编码为视频（VCD）。使用可变比特率编码，最大比特率为2.6Mbit/s。IEC 62107定义了超级VCD标准。

XSVCD尽管不是工业标准，但增加了视频的分辨率和比特率，从而改善了SVCD上的视频质量。XSVCD仍然采用MPEG-2视频，分辨率为720×480（29.97Hz的帧率）或720×576（25Hz的帧率）。仍然使用可变比特率编码，最大比特率为9.8Mbit/s。

S-Video 分离视频（Separate Video），也称为Y/C视频。使用的是分离亮度（Y）和色度（C）视频信号，而不是单一的复合视频信号。通过简单地将Y与C信号相加，就可以产生复合视频信号。

当使用Letterbox模式的画面格式时，在C信号上可能出现+2.3V的直流分量偏移。当使用16：9的变形画面格式时，在C信号上可能出现+5V的直流分量偏移。标准的4：3接收设备忽略了所有的直流分量偏移，这样显示的是典型的Letterbox模式画面。

IEC 60933-5标准规定了S-Video连接头，包括信号电平。

SVM 参见速度扫描调制（Velocity Scan Modulation）。

同步（Sync） 对于显示任何类型的视频，同步是基本的信息，必须得到同步信息。本质上讲，同步信号告诉显示器在什么地方显示图像。线同步（HSYNC）告诉显示器将图像放置在左右方向上的什么地方。场同步（VSYNC）告诉显示器将图像放置在上下方向上的什么地方。

模拟SDTV和EDTV信号使用双极性电平同步，其中同步电平为低于消隐电平的已知值。模拟HDTV信号使用三极性同步，其中同步电平为高于和低于消隐电平的已知值。

模拟HDTV使用三极性同步信号是为了时序的准确度。双极性同步信号的水平时序参考点定义于线同步脉冲前沿50%的点。为了精确地确定这个点，必须确定消隐电平和同步提示电平，并确定它们的中点值。如果该信号存在任何的失真，那么都将降低时序准确度。

对于三极性同步信号，时序参考点是同步信号的上升沿与消隐电平的交点。准确确定该点要容易得多，并且能够相对容易地实现。它也更不易受信号失真的影响。

同步生成器（Sync Generator） 同步生成器是一种产生同步信号的电路。同步生成器可以具有同步锁相（genlock）功能。

同步噪声门（Sync Noise Gate） 同步噪声门用于在模拟视频信号波形中定义一个区域，视频解码器在这个区域中查找同步脉冲。在该定义窗口之外的所有信号将被忽略。同步噪声门的主要目的是确保视频解码器的输出是完好、干净和正确的。

同步剥离器（Sync Stripper） 模拟视频信号包含视频信息（即用于在显示器上显示的画面）和时序（同步）信息。时序信息告诉接收设备在显示器的什么地方显示信息。同步剥离器就

是要将同步信息从模拟视频信号中剥离出来，并丢弃其他信息。

同步（Synchronous） 指的是在一个系统中，两个或多个事件同时发生。

SVCD 参见超级VCD（Super VideoCD）。

图文电视（Teletext） 与视频信号一起传输数据的一种方法。ITU-R BT.653列出了世界上使用的主要图文电视系统。而ETSI ETS 300 706详细地定义了PAL的图文电视标准。北美广播图文电视规范（NABTS）为525线的C系统。

对于数字传输，如HDTV和SDTV，图文电视字符是作为一个独立的流与音视频数据一起以复用方式进行传输的。实际上，通常的做法是将这些流嵌入到MPEG视频比特流本身中，而不是嵌入到传输层。遗憾的是，这样的图文电视流没有宽屏标准，每个系统（DSS、DVB、ATSC、DVD）都有自己的解决方案。

细分同步（Tessellated Sync） 这就是欧洲所谓的锯齿同步（serrated sync）。参见锯齿脉冲和同步（Serration Pulse and Sycn）。

时基校正器（Timebase Corrector） 某些视频源打乱了视频的同步信号。一个最常见的视频源就是VCR。时基校正器就是要对同步时序不好的视频信号进行调整。

三极性同步（Tri-Level Sync） 具有三种电平的同步信号，通常用于模拟HDTV信号。参见同步（Sync）的定义。

真彩色（True Color） 真彩色表示图像的每个样本点都是单独地用三个颜色分量来表示，如R'G'B'或YCbCr。

欠扫描（Underscan） 当对一幅图像进行显示时，如果图像的所有部分在显示器上都是可见的，包括图像的上下左右边，那么就说该图像是欠扫描的。在计算机显示器中的图像通常是欠扫描的。

上行链路（Uplink） 在地面基站上用于将信息发送到卫星的载波。

V型芯片（V-chip） 参见CEA-608。

可变比特率（Variable Bit Rate） 可变比特率（VBR）是指比特流（压缩的或未压缩的）每秒中传输的数据比特数是变化的。对于简单的场景可以赋予低比特率，对于复杂场景则使用较高的比特率。这使得可以在更一致的水平上保持音视频质量。

VBI 参见场消隐期（Vertical Blanking Interval）。

VBR 可变比特率（Variable Bit Rate）的缩写。

VC-1 SMPTE 421M视频编解码原来的名称，它是基于微软的WMV9 视频编解码的。

VCD 为VideoCD的缩写。

速度扫描调制（Velocity Scan Modulation） 通常用于基于CRT的电视机，以增加图像的外显锐度（apparent sharpness）。在水平方向上由暗到亮的转变，光束的扫描速度立即增加到逼近转变的颜色，这使得显示器在亮度变换之前相对较黑。一进入较亮的区域，光束的速度立即减慢，使得在变换之后很短时间内，显示器相对更亮。当由亮转变为暗时，这个过程刚好相反。

场消隐期（Vertical Blanking Interval，VBI） 在场消隐期内，视频信号电平等于消隐电平，使得在由CRT显示屏底部回扫到顶部时，不显示电子束。

场期时间码（Vertical Interval Timecode） 在每个场消隐期内，时间码信息存储在扫描线中。

垂直分辨率（Vertical Resolution） 参见分辨率（Resolution）。

垂直扫描速率（Vertical Scan Rate） 对于非隔行视频，垂直扫描速率与帧率是一样的。对于隔行视频，垂直扫描速率通常为场速率的一半。

场同步（Vertical Sync） 场同步是模拟视频信号的一部分，它告诉视频解码器从显示器的什么地方开始显示图像的顶部。

残留边带（Vestigial Sideband） 这是一种将数字数据编码到载波上用于射频发射的方法。在美国，8-VSB用于在空中广播ATSC HDTV信号。

视频载波（Video Carrier） 这是一个特定的频率，在与音频数据混合和发射之前先调制视频数据。

视频接口端口（Vieo Interface Port） 这是一种数字视频接口，设计这种视频接口是为了简化视频IC卡的相互连接。其中一部分是数字视频接口（基于BT.656），是用于简化视频IC卡的相互连接。第二部分是主处理器接口（host processor interface）。VIP（Vieo Interface Port）是一个VESA规范。

视频混合（Video Mixing） 视频混合采用两个独立的视频源（它们必须是同步锁定的），并将它们合并在一起。参见alpha混合（Alpha Mix）。

视频调制（Video Modulaion） 将基带视频信号转换为射频（RF）信号。

视频模块接口（Video Module Interface, VMI） 这是另外一种数字视频接口，是设计用于简化视频IC卡之间的互连的。

视频点播（Video-On-Demand） 视频点播（Video-On-Demand或VOD）准许用户任意选择观看哪个节目，在选择确定以后，几乎立即就开始播放。当用在因特网或其他网络上时，通常称为“流视频（Streaming Video）”。对于广播、卫星和有线网络，通常称为“计次付费（pay-per-view）”，通常限于特定的时间开始播放。正因如此，也可以称之为“近似视频点播（near-video-on-demand）”或NVOD。

视频节目系统（Video Program System，VPS） VPS在某些国家用于代替PDC来控制VCR。数据格式与PDC是一样的，但VPS是场消隐期内在特定的扫描线上传输，通常使用的是线16。

VCD（VideoCD） 是一种紧凑磁盘，可以存放最多大约1小时的数字音视频信息。使用MPEG-1视频，分辨率为352×240（帧率为29.97Hz）或352×288（帧率为25Hz），音频使用MPEG层2，具有固定比特率244kbit/s，支持两个单声道或一个立体声通道（具有可选的Dolby®专业逻辑）。使用固定比特率（1.15Mbit/s）编码。下一代VCD在中国定义为超级VCD。

尽管XVCD不是企业标准，但是它提高了视频的分辨率和比特率，从而改善了VCD的视频质量。使用的仍然是MPEG-1视频，分辨率可达720×480（29.97Hz的帧率）或720×576（25Hz的帧率）。仍然使用固定比特率编码，比特率为3.5Mbit/s。

VIP 参见视频接口端口（Video Interface Port）。

VITC 参见场消隐期时间码（Vertical Interval Timecode）。

VMI 参见视频模块接口（Video Module Interface）。

VOB 使用VOB文件在DVD上存储的DVD视频电影。VOB视频电影通常包含复用的杜比数字音频和MPEG-2视频。VOB文件按如下方式命名：vts_XX_Y.vob，其中XX表示标题，Y为标题的部分（part）。可以有99个标题和10个部分，尽管vts_XX_0.vob不包含视频，但通常包含节目单或导航信息。

VOD 参见视频点播（Video-On-Demand）。

VPS 参见视频节目系统（Video Program System）。

VSB 参见残留边带（Vestigial Sideband）。

VSM 参见速度扫描调制（Velocity Scan Modulation）。

VSYNC 请查看场同步（Vertical Sync）的定义。

白色电平（White Level） 对于特定的视频系统，该电平定义了什么样的颜色是白色。

宽屏信令（Wide Screen Signaling，WSS） WSS可以用于576i的线23和480i的线20和283，用于指定节目的宽高比和其他信息。16：9的电视机可以使用该信息，使得可以按正确的宽高比显示节目。ITU-R BT.1119和ETSI EN 300 294规定了576i和480i系统的WSS信令。EIA-J CPR-1204和IEC 61880也为480i系统规定了另外一种WSS信令。

世界图文电视广播系统（World System Teletext，WST） BT.653 525线和625线系统的B类图文电视。

WSS 参见宽屏信令（Wide Screen Signaling）。

WST 参见世界图文电视广播系统（World System Teletext）。

XSVCD 为扩展超级VCD（eXtended Super VideoCD）的缩写。参见超级VCD（Super VideoCD）。

XVCD 参见扩展VCD（eXtended VideoCD）的缩写。参见VCD（VideoCD）。

xvYCC xvYCC（扩展色域YCbCr视频）颜色空间对一般的YCbCr的色域进行了扩展，使得可以再生出比YCbCr多出1.8倍的颜色。xvYCC（IEC 61966-2-4）规范使用BT.709色度和D65参考白色。scR′G′B′与xvYcbCr之间的变换公式和R′G′B′与YcbCr之间的变换公式相同。

基于xvYcbCr的YCbCr数据有8位，表示范围为1~254，使得可以与现有设计向后兼容。Y有8位，表示范围为－15/219~＋238/219（－0.068493~＋1.086758）；CbCr数据有8位，表示范围为－15/244~＋238/224（－0.066964~＋1.062500）。

Y/C视频（Y/C Video） 参见S视频（S-Video）。

Y/C分离器（Y/C Separator） 在NTSC和PAL系统中，Y/C分离器是用于在NTSC和PAL视频解码器中将亮度和色度信息分开的电路。这是任何NTSC/PAL视频解码器必须做的第一件事情。复合视频信号进入Y/C分离器，然后进一步解码就可以得到色度。

YCbCr YCbCr是起初在BT.601定义的颜色空间，现在用于所有的数字分量视频格式。Y为亮度分量，Cb和Cr分量为色差信号。技术上正确的表示应该是Y′Cb′Cr′，因为所有的三个分量都是由R′G′B′推导而来。大多数人使用YCbCr表示而不是Y′CbCr或Y′Cb′Cr′表示。

4：4：4的YCbCr表示每个Y采样，有一个Cb和Cr采样。

4：2：2的YCbCr表示每两个水平的Y采样，有一个Cb和一个Cr采样。

4：1：1的YCbCr表示每四个水平的Y采样，有一个Cb和一个Cr采样。

4：2：0的YCbCr表示每个2×2的Y样本块，有一个Cb和一个Cr采样。4：2：0的YCbCr有三个变种，差异在于Cb和Cr相对于Y的采样位置。

注意，对于SDTV和HDTV应用，将R′G′B′转换为YCbCr所使用的参数是不一样的。

YIQ YIQ是NTSC视频系统使用的可选颜色空间。Y分量为图像的黑白部分。I和Q部分是色微分量；实际上它们只是放在黑、白和亮度上的颜色分量。很多人使用YIQ表示，而不是Y′IQ或Y′I′Q′。在技术上正确的表示应该是Y′I′Q′，因为它们都是由R′G′B′推导而来的三个颜色分量。

YPbPr YPbPr为YCbCr颜色空间的模拟版本，它具有特定的电平和时序信号，设计YPbPr是为了与设备的接口相连。消费类视频标准由EIA-770定义，专业视频标准由大量的SMPTE标准定义。EIA-770使用的VBI数据格式由CEA-805定义。很多人使用YPbPr表示，而不使用Y′PbPr或Y′Pb′Pr′表示。在技术上，正确的表示应该是Y′Pb′Pr′，因为其所有三个分量都是由R′G′B′推导而来的。

YUV YUV是由NTSC和PAL视频系统使用的颜色空间。与UIQ颜色空间一样，Y为亮度分

量，而U和V为色微分量。很多人使用YUV表示，但它们实际表示的是YCbCr数据。大多数人使用YUV表示，而不使用Y'UV或Y'U'V'。在技术上，正确的表示应该是Y'U'V'，因为其所有的三个颜色分量都是由R'G'B'推导而来的。

YUV也是某些消费类设备的分量模拟接口的名称。有些制造厂商不正确地将其标注为YCbCr。THX认证要求将YUV标注为YPbPr。

YUV9 Intel公司的4：1：0 YCbCr格式。画面被分为块，每个块由4×4的样本组成。对于每个块，赋予了16个8位的Y值、一个8位的Cb值和一个8位的Cr值。其结果为平均每个像素有9位数据表示。

YUV12 Intel公司对以平面格式（planar format）存储于内存中的MPEG-1 4：2：0 YCbCr数据的一种表示。画面被分为块，每个块由2×2的样本组成。对于每个块，赋予了4个8位的Y值、一个8位的Cb值和一个8位的Cr值。其结果为平均每个像素由12位数据表示。

YUY2 Intel公司对4：2：2 YCbCr格式的表示。

拉链纹伪影（Zipper） 参见拉链伪影（Creepy-Crawlies）的定义。

放大（Zoom） 放大（zoom）是图像伸缩（scaling）的一种类型。放大就是使图像变大，从而可以看到更多的细节。在伸缩（scaling）定义中描述的例子也可以作为此处定义的例子。

双声道模式（Zweiton） NTSC和PAL视频中用于实现立体声或双－单声道音频的技术。第一个FM副载波传输L＋R信号，第二个FM副载波传输R信号（对于立体声）或第二个L＋R信号。在BS.707中对此进行了讨论，这是类似于BTSC的一种技术。



索　引

索引中的页码为英文原书页码，与本书中页边标注的页码一致。

数字

A

B

D

E

F

G

H

I

J

K

L

M

N

O

P

T

U

V

W

X

Y

Z



Video Demystified
(Fifth Edition)

视频技术手册（第5版）

“在视频技术处理方面，这是目前堪称‘圣经’的参考指南。”

——**Amazon.com**

“视频技术涉及的内容非常丰富，难得看到像本书这么全面并且实用的手册，你还犹豫什么呢？绝对物超所值！”

——**Bill Schweber, *Planet Analog*资深编辑**

视频技术的应用和开发是目前信息技术领域最热门的话题之一，3G通信、高清晰度电视、数字电视和宽带网络等都是围绕如何有效处理和传输视频信息而展开研究的。本书回答了全世界的工程师提出的各种视频相关的问题，自第1版问世以来，读者好评如潮，已经成为业内视频技术处理的经典参考指南。

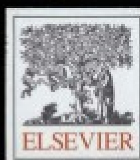
本书涵盖了视频技术领域的最新成果和未来发展方向，系统讲述视频技术的基础理论和实践技能，重点介绍最新的视频编解码器（如H.264、VC-1、Dolby Digital® Plus、DTS-HD等）的实现原理，同时详尽阐述如何解决目前最先进的消费类电子设备（如HD DVD、数字机顶盒、数字电视和网络电视机顶盒等）中所涉及的视频编解码问题，并概述隐藏字幕、图文电视、V-chip技术、复杂的图像合成技术、基于IP网络的ARIB与DVB技术以及DRM的最新技术发展情况。书中还列出了最新的广播、电缆和卫星规范。可以说，有本书在手，各种视频系统处理方面的问题都可以找到标准答案。

Keith Jack 资深电子工程师和技术经理。现任知名无晶圆厂半导体供应商Sigma Designs公司产品销售主管，之前曾在哈里斯半导体、罗克韦尔等公司担任芯片设计工作，参与设计过40余款消费类电子产品的多媒体芯片，有丰富的视频技术经验。

本书译自原版*Video Demystified (Fifth Edition)*，并由Elsevier授权出版。









ISBN 978-7-115-21020-3/TN
定价：119.00元

本书相关信息请访问：图灵网站 http://www.turingbook.com
读者/作者热线：(010) 51095186
反馈/投稿/推荐信箱：contact@turingbook.com

分类建议：电子电气／视频处理
人民邮电出版社网址 www.ptpress.com.cn